PHYSIOLOGICAL REVIEWS

VOL. XIII JULY, 1933 No. 3

A REVIEW OF FLUORINE AND ITS PHYSIOLOGICAL EFFECTS

F. J. McCLURE

Institute of Animal Nutrition, Pennsylvania State College

As early as 1670 it was known that glass is attacked by the fumes produced when fluorspar is treated with sulfuric acid. It was not until 1802, however, that Morichini used this reaction, apparently for the first time, to test for the presence of these same "fumes" in a substance of animal origin. As a result of his studies he reported the presence of F in fossil teeth. Later (1805) Gay-Lussac and Berthollet found F in enamel of normal teeth and thus opened the question, which continues to this day, as to the extent and significance of the occurrence of F in teeth and other body substances. Some time later F was found present in the ash of blood and of milk; in 1869 it was reported present in the brain; and in 1888 its discovery in the yolk and shell of the egg brought forth the suggestion of an important physiological rôle. In 1912 Gautier and Clausmann, who have published the most extensive analytical results for F, presented their first work and have since reported results which, together with those of other investigators indicate that F is widely distributed in soils and waters and in plant and animal substances.

The toxicity of fluorides to man became evident when the chemists Thenard and Davy, while attempting to prepare pure F, were made seriously ill by accidentally breathing the vapor of hydrofluoric acid, and when Louyet and Nickles lost their lives from the same cause. Such violent effects naturally occasioned concern as to the probable toxicity of F contained in drinking waters, in wines and beers, in phosphate baking powders, and in other foodstuffs in which F might occur naturally or, as was the case following the discovery of its antiseptic properties, might be present in the form of a preservative. Experimental evidence of the toxic effect of F on man and animals is, therefore

abundant in the literature. Interest in the physiological rôle of F was stimulated somewhat by the observation made by McCollum et al., on the effect of F on the teeth and by the work of Embden and his school as to its effect on muscle metabolism, an effect not unrelated to its action on certain enzymatic changes. A group of investigators in this country have studied the effects of ingested fluorides for the sake of assessing the significance of the small amount of F which occurs naturally in rock phosphate and thus becomes a not inconsiderable part of the intake of animals fed rock phosphate as a calcium mineral supplement. During the past year evidence has been presented which indicates that F in drinking water may be the cause of a defect of human teeth known as mottled enamel. Most experiments planned, therefore, to study the physiological effects of F have had to do with the following problems: *a*, the distribution and purpose of F in the animal body; *b*, the effect of small quantities of ingested fluorides on food consumption, growth and reproduction; *c*, the relation of F to the development of teeth and bones, and *d*, the effect of fluorides on certain enzymatic changes related particularly to fat and carbohydrate metabolism.

DISTRIBUTION OF FLUORINE. Disagreement among analysts reporting quantitative results for F, wherever it occurs in very small amounts, may be attributed largely to difficulties involved in the analytical procedure. The notable chemical activity of F, as well as the volatile nature of many of its compounds, has contributed to these difficulties. Despite the fact that numerous methods have been devised and have been used, at least by their originators, it is not established that any F method now known is adequate for the quantitative analysis of plant and animal substances. The most promising methods for the purpose, however, are *a*, colorimetric methods (1) (2) (3) (4) (5); *b*, gravimetric methods based on the formation of CaF_2 (6) (7) and on the formation of fluorides of rare metals, as thorium fluoride ($ThF_4 \cdot 4H_2O$) (8), and lanthanum fluoride (LaF_3), (9), or *c*, methods which consist in passing volatilized SiF_4 into water and then determining, either volumetrically or gravimetrically, the amount of H_2SiF_6 formed (10) (11) (12) (13) (14).

a. In minerals, soils, and waters. Fluorine occurs most abundantly in nature in the mineral fluorspar (CaF_2) and in cryolite, a double fluoride of aluminum and sodium. Minerals containing small amounts of F as biotite, tourmaline, muscovite, and apatite, distributed throughout the earth's crust, account for part of the F found present in soils. According to Steinkoenig (15) F in soils amounts to about 0.03 per cent.

Jacob and Reynolds (16) have investigated the F content of various commercial grades and types of phosphate rock. The amounts of F found vary from 2.62 per cent in Tennessee phosphate "run of mine material" to 4.08 per cent for "carefully selected lump rock." A "fertilizer grade" of the same rock contained 3.72 per cent F, while a high grade washed and ground rock contained 3.89 per cent. These values are similar to those obtained for samples of phosphate rock occurring in North Africa, Idaho, Wyoming, and Florida.

The presence of F in mineral, river and sea waters seems established. Carles (17) found mineral waters containing 0.001–0.018 gram F per liter. River water analyzed by Gautier and Clausmann (18) contained 0.02 to 0.60 mgm. F per liter. Carnot (19) found 0.822 mgm. F per liter of ocean water. Smith, Lantz and Smith (20) found that 3.8 to 7.15 mgm. F per liter in drinking water caused mottled teeth in humans. Other waters not harmful showed 0.0–0.3 mgm. F per liter. Churchill (21) reports from 2.0 to 13.7 mgm. F per liter in drinking waters of localities where mottled enamel is endemic.

b. In plants and other foodstuffs. Gautier and Clausmann (22), on examination of 63 plant substances, found an average of 2.65 mgm. (0.00265 per cent) F per 100 grams of fresh material. Apparently more F is present in leaves than in stems and there is more in the skin of fruit than in the pulp. According to these authors the F content of some edible substances (reported on the fresh basis) is as follows: corn meal 0.00117 per cent; rye meal 0.00052 per cent; kidney bean 0.00170 per cent; cabbage 0.000088 per cent; cauliflower 0.00021 per cent; apple pulp 0.000034 per cent; apple skin 0.00076 per cent; strawberry 0.00012 per cent; potato 0.000084 per cent; tomato 0.00020 per cent, and buckwheat 0.00217 per cent. The average for 32 representative edible food substances is 0.000573 per cent F for the fresh material.

Woodman and Talbot (23) found "that the presence of a slight amount of F in malt liquors is the rule, rather than the exception," and "that the principal source of this fluoride is the malt itself." They set a maximum limit of permissible fluoride in malt liquors at 10 mgm. per liter.

c. In bones and teeth. In 1802 Morichini (cited by Gay-Lussac and Berthollet) (24) found F in fossil teeth and later (1805) he, as well as Gay-Lussac and Berthollet (24), was convinced of the presence of F in normal human teeth. In 1807 Berzelius (25) reported 1.0–2.0 per cent F in bones but later found his results to be erroneous. It was believed generally, during the time of Berzelius, that fresh bone contains much

less F than fossil bone. However, Berzelius believed F not to be an essential constituent of normal bone.

Owing to the difficulties involved in the direct determination of F it became more and more the custom to report the F in bone by difference and thus Weiske (26) reported 3.8 per cent F in the ash of chicken bone; and results as high as 5.0–8.0 per cent F in the whole bone were recorded. The values obtained by Carnot (27), varying from 0.1–0.3 per cent for bone ash, agree with those established by Zalesky (28).

In experiments conducted by Brandl and Tappeiner (29), the deposition of F compounds in the body following F feeding was demonstrated, but no F was found in the ash of normal bones. Gabriel (30) made a thorough study of the method of Fresenius, modified the method somewhat, and by means of numerous careful analyses of teeth and bones from various sources obtained the low value of 0.05 per cent as the F content of normal tooth and bone ash. In a few samples of ox bone ash he found 0.10 per cent F. Gabriel thought that his results might be taken as indicating that bones and teeth are entirely free of F.

Wrampelmeyer (31) found from 0.65 per cent to 1.40 per cent F respectively in sound and carious teeth of children and adults. Having found F in bone ash in amounts varying between 0.05 and 0.32 per cent, Jodlbauer (32) regarded F as a constant constituent of bone tissue.

However, the gas volume method used by Jodlbauer has been justly criticized by Gassman (33), who showed that warming bone ash with concentrated H_2SO_4 formed gaseous HCl as well as SiF_4, and that consequently volumetric measurements of gases thus formed did not indicate SiF_4 quantitatively. Gassman believed that normal bones do not contain F.

Among the more recent reports of the F content of bones and teeth there should be mentioned the results of Gautier and Clausmann (34) (35), Trebitsch (36), and Bethke, Kick, Edgington and Wilder (37). The results of Gautier and Clausmann for teeth and bones show the approximate values of 150 mgm. of F per 100 grams of dry dental enamel and 46 mgm. per 100 grams bone ash. The Meyer and Schulz (9) method was used by Trebitsch. His values for F in the dry substance of human bones are extremely high (0.29–0.59 per cent). According to Bethke, et al., the ash of the femurs from two lots of pigs without added F in their ration contained 0.0409 per cent and 0.0231 per cent F respectively; representing 0.0251 per cent and 0.0142 per cent of F respectively for whole dry bone.

In summarizing the results which have been reported for bones and

teeth, it is to be noted that Gautier and Clausmann obtained values for bone ash varying between 17.5 mgm. and 123.0 mgm. F per 100 grams of ash. Jodlbauer's results are high, due, no doubt, to errors in his method. His values vary between 50 mgm. and 540 mgm. F per 100 grams of ash, and average 191 mgm. The values of Zalesky and Carnot are between 170 mgm. and 310 mgm. F per 100 grams of ash. Since normal bone contains approximately 60 per cent of ash, the average values of Gautier and Clausmann represent about 0.0298 per cent F in the whole bone. Bethke et al. report 0.0251 per cent and 0.0142 per cent for normal bone. The majority of the analytical results appearing in the literature may be regarded as indicating that normal bone contains between 0.01 per cent and 0.03 per cent F, though there is evidence that great variations in F content occur. It is to be recalled that Gabriel, Brandl and Tappeiner, as well as Gassman, believed bones not to contain F.

The results of F determinations on teeth are more limited and perhaps more variable than those reported for bones. Gautier and Clausmann's results vary from 0.118 per cent to 0.180 per cent F in dental enamel on dry basis. Carnot found dental enamel on the dry basis to contain 0.10 per cent–0.30 per cent F. It cannot be said that it is established that the F content of teeth is as high as these values indicate. According to some investigators there is less F in teeth than in bones, or as reported by Casares and Casares (59), who have recently perfected a qualitative test for F in bones and teeth, F does not exist in teeth. However, the quantitative results for F which appear in the literature up to the present time indicate the presence of about 0.10–0.20 per cent F in normal dental enamel.

d. In other animal organs and tissues. Of the tissues examined by Gautier and Clausmann (38), the richest in F are the following: dental enamel containing 118 mgm. to 180 mgm. of F (these values and those immediately following represent milligrams of F per 100 grams of dry substance); bone (diaphysis) 56.0 mgm. to 87.0 mgm.; epidermis 16.4 mgm.; hair 6.1 mgm. to 19.7 mgm.; thymus 3.9 mgm. to 11.1 mgm.; testicle 1.2 mgm. to 4.2 mgm.; blood 2.3 mgm. to 4.4 mgm.; and hair 1.0 to 3.0 mgm. Tissues poorest in F are cartilage, 0.4 mgm. to 1.4 mgm., and muscle, 0.04 mgm. to 0.46 mgm.

Apparently blood is high in F (38) (39) (40). According to Gautier and Clausmann (38), 100 grams of dry blood contain approximately 3.5 mgm. of F, depending somewhat on the species. Brandl and Tappeiner (29) found 120 mgm. NaF (54.27 mgm. F) in 100 grams dry blood and

19 mgm. NaF (8.59 mgm. F) in 100 grams of fresh blood. Stuber and Lang (41) found values varying from 0.0 mgm. to 0.94 mgm. F per 100 grams of fresh blood. In most cases the results which they obtained were thought to be abnormally high, due presumably to the presence of relatively large amounts of F in the drinking water.

Three samples of fresh cow's milk, as analyzed by Stuber and Lang, contained respectively 0.11 mgm., 0.16 mgm. and 0.20 mgm. F per 100 grams. Gautier and Clausmann report 0.046 mgm. and 0.051 mgm. F for two samples of human milk, and 0.180 mgm. F for cow's milk, per 100 grams of fresh milk.

SIGNIFICANCE OF THE OCCURRENCE OF FLUORINE IN ANIMAL ORGANS AND TISSUES. It is apparent that F is found in many of the organs and tissues of the body and may be excreted from the body in both the urine and feces. Schulz (42) first expressed the opinion that F is excreted through the urine combined with alkali. Tappeiner (43) fed a dog 0.5 gram NaF, and the urine of the two following days contained 0.124 gram of NaF. Another dog received 0.240 gram of NaF; 0.059 gram appearing in the urine of the two following days. Brandl and Tappeiner (29) fed varying daily amounts of F to a young growing dog for a period of about 21 months. The total NaF in the feed for the period amounted to 402.9 grams and in the feces and urine it amounted to a total of 330.5 grams; 72.6 grams of NaF apparently were deposited in the organism. The blood, muscles, liver, skin, bones, cartilages, and teeth were weighed and analyzed and a total of 64.64 grams of NaF reported present in the whole body. Gautier and Clausmann (38) estimate a daily excretion by man of 0.23 mgm. F in the urine, and 0.80 mgm. in the feces (a total of 1.03 mgm.). F in the diet probably exceeds 1.0 mgm. per day. To account for further loss of F from the body Gautier and Clausmann, as well as Brandl and Tappeiner, call attention to the accumulation of F in the hair, nails, and epidermis as a means of F excretion. It may be concluded from the foregoing facts that the metabolism of F is probably a normal function of the animal organism.

Gautier and Clausmann have been led to believe that F plays a "necessary and a specific rôle" in the animal body. Gautier (44) classifies the different tissues of the body into three groups, according to their F content. The first group is comprised of organs of assimilation and secretion, highly differentiated and having an active metabolism, but containing very small amounts of F. A second group of tissues comprises the bones, cartilage, tendons, etc., which contain F varying from 88.0 mgm. (bone) to 4.5 mgm. (cartilage) per 100 grams of dry matter.

The third group comprises organs for "mechanical" protection or defense or for ornamentation, such as hair, feathers, nails, epidermis, etc. These tissues are rich in F, containing about 180 mgm. per 100 grams of dry tissue. Fluorine in these tissues he regards as playing a secondary rôle associated with the hardness, the resistance, and to a certain degree, "*d'inalterabilite chemique*" of such tissue.

Calcium fluophosphates occur naturally and may represent a form in which F could occur in bones and teeth, since the formula for the calcium-phosphorus compound of bones and teeth may not be known (45). According to Trebitsch (36) it has been suggested, as a result of Roentgen-analysis, that the hardest apatite, namely, fluor-apatite, $3[Ca_3(PO_4)_2]\cdot CaF_2$ occurs in dental enamel and that the hardness of teeth depends on the presence of crystal masses of apatite. Thus there is a tendency to regard F as an essential constituent particularly of the teeth. Recently there has been presented clinical as well as experimental evidence of the fact that mottled enamel, a defect of human teeth, may be an endemic imperfection caused by the action of F present in the drinking water. These and other experiments have demonstrated the fact that, regardless of their F content, teeth are influenced structurally and presumably chemically by the presence of extremely small amounts of F in the diet. For this reason it appears that the intake of the element must be kept at an extremely low level. On the other hand no experimental evidence has demonstrated the indispensability of F or the specific effects of the complete withdrawal of F from the diet. Mere traces of F are known to exert a potent influence on the reactions of certain enzymes (as well as on tooth formation) and further investigation may reveal a specific rôle of F in certain chemical changes necessary for normal cell metabolism.

PHYSIOLOGICAL EFFECTS OF FLUORINE. *a. Fluorine toxic to man and other animals.* Numerous cases are on record which show that the ingestion of F compounds beyond certain limits is fatal to man and other animals (46) (47). A clinical picture of the effect of the ingestion of small doses of fluorides is described by Tappeiner (43) and by Schulz (42). Tappeiner used dogs, rabbits, guinea pigs and cats as experimental animals. As high as 0.5 gram NaF per 100 grams body weight was given internally and 0.15 gram was given by injection subcutaneously and intravenously. The following characteristic symptoms were observed: 1, a condition of drowsiness and weakness resulting from paralysis of the vaso-motor centers; 2, cramps which may attack a single organ or the entire body, and are epileptic in character. They

are not so evident in cases of internal poisoning, although in many instances they dominate the poisoning effects; as non-reflex changes, they are independent of the disturbance of the breathing and circulation and proceed from the spinal cord and more highly centralized organs; 3, paralysis of the vasomotor centers; 4, acceleration and deepening of the breathing with paralysis following; 5, vomiting; 6, secretion of the salivary and tear glands, which is not stopped with atropin; 7, early rigor following death.

There is unmistakable evidence in the data reported by Goldemberg (48) that the injection of NaF lowers the basal metabolic rate of white rats. Goldemberg employed the method devised by Haldane for the indirect determination of heat production. The heat produced by the experimental animal was determined before and after the intraperitoneal injection of the fluoride solutions. Amounts of NaF varying from 15 to 18 mgm. per kgm. of body weight lowered the heat production per kilogram per hour as much as 12 per cent to 63 per cent below the previous rate. The effects were noted usually 15 minutes to about 60 minutes after the injection.

Goldemberg (49) also observed a local ulcerous condition at the point of injection of fluorides, as well as toxic symptoms similar to those reported by Tappeiner (43) and Schulz (42).

The work of Cristiani and Gautier (50) (51) (52) (53), of Cristiani and Chausse (54) (55), and Cristiani (56) (57) has generally confirmed the results of earlier experiments as regards the symptoms of F poisoning. This group of investigators believe there is an absorption of F by the vegetation in the vicinity of aluminum plants which give off waste gases containing F. Apparently the herbivorous animals of the vicinity are in some way affected by the vegetation which they consume. Cristiani and Gautier (51) claim to have reproduced toxic effects characteristic of fluorides by feeding the suspected forage to guinea pigs. Sette (58) regards these toxic symptoms as a general acid intoxication, which might result from HNO_3 and H_2SO_4 vapors contained in the air.

It becomes obvious on inspection of the data on the subject (61) (62) (63) (64) (65) (66) (67) (68) (69) that the form of F used, the method and length of time of administration, as well as individual susceptibilities, have varied the toxic effects to such an extent that a comparison of the data obtained is scarcely warranted. However, there is evidence that fluosilicates are more toxic than either NaF or CaF_2 and that NaF is more toxic than CaF_2. Other constituents of the ration may perhaps

have affected the results in numerous instances. The level of toxicity of F in any form, for man particularly, is of some importance because of the fact that the antiseptic property of fluorides has occasioned their use as a preservative in wines and beers, tomato catsups, various plant juices, etc. There is still some controversy as to the desirability of their use for this purpose and it would appear that more work is needed to determine at what level F may be used in this connection without its becoming toxic. As was mentioned, drinking waters containing quantities of F varying from 3.8 to 7.15 mgm. per liter may affect detrimentally normal tooth development and may, therefore become the source of a F intake which exceeds the hygienic level.

b. *Effect of fluorine on development of bones and teeth: Effect on amount of fluorine present.* There is evidence that the F of bones and teeth, as well as of other body organs and tissues, may be increased by an increase in the F intake. Sonntag (59) found that, whereas normal teeth and bones contain about 0.3 per cent F, the teeth of animals fed NaF contain 1.73 per cent F and the bones 1.29 per cent F. Christiani and Gautier (50) (51) (52) believe that F tends to accumulate in the bony tissue producing a condition resembling osteomalacia. According to Cristiani (56) the deposition of F in bone is a very reliable index of F toxicity.

Bethke et al. (37) found that the ash of femurs of two groups of pigs receiving limestone or steamed bone meal as a Ca mineral supplement contained 0.0409 and 0.0231 per cent F respectively, while bone ash of femurs from lots of pigs receiving varying amounts of F in the form of NaF or rock phosphate contained 0.528 per cent, 0.7935 per cent, 1.1077 per cent, 0.5937 per cent, 1.0934 per cent, and 1.0369 per cent F.

The bones of a young growing dog fed NaF in small amounts by Brandl and Tappeiner (29) for a period of about 21 months were examined microscopically. These investigators found deposited in the bones, vivid, glistening crystals which they regarded as crystalline compound similar to fluorite (CaF_2). They believed that soluble fluoride salts in the feed are deposited in the bones in considerable amounts but they were unable to find any crystals in the teeth. An increased brittleness of some of the bones was noted also.

Structural changes. Structural changes in teeth, following the ingestion of small amounts of F, were observed by McCollum, Simmonds, Becker and Bunting (65). Sodium fluoride to the extent of 0.0226 per cent (0.0101 per cent F) was included in an otherwise normal diet of rats. It was observed that the rats receiving the added F developed

incisors which were abnormal in color, lacking the characteristic orange tint on the anterior surface. There were striking changes in the shape of the teeth. The upper incisors grew backward into a circle, while the corresponding lower incisors were shortened.

These changes appear to be due to an alteration in the structure and hardness of the teeth which, according to McCollum et al., resulted finally in "the lack of the natural wear which the incisors of the rat under normal conditions exert on each other in occlusion." It is known that if for any reason this incisal reduction by wear does not take place, the incisors undergo a marked over-growth quite similar to the over-growth occurring in the teeth of rats fed F. Apparently the lower incisors, due to their inferior quality, are promptly fractured while the upper incisors, lacking the necessary wear, grow to an abnormal length. The authors state that the manifest inferior quality of these teeth indicates "a retrograde disturbance of tooth development, rather than a stimulation to overactivity." Chaneles (70) made a microscopic study of the teeth of rats fed fluorides and describes, in detail, changes in the structure and arrangement of the cells.

Changes similar to those observed by McCollum in teeth of rats were produced by Tolle and Maynard (72) by feeding rock phosphate, and by investigators at the Arkansas Agricultural Experiment Station (73), who fed NaF and CaF_2. Bergara (74) observed such changes by feeding NaF, and stated also that the incisor teeth lost their reddish yellow color, finally showing dark bands. At first the teeth seemed more opaque and lighter.

Results obtained at the Michigan Agricultural Experiment Station (75) (76) have demonstrated an effect of F on teeth of dairy cattle by the feeding of either $Na_2SiF_6 \cdot 2H_2O$ or rock phosphate. The mandibles of these cattle were much thicker and the surfaces were rougher than those of normal cattle. McClure and Mitchell (68) again demonstrated that "a level of 0.03 per cent and 0.06 per cent fluorine in the form of sodium fluoride or calcium fluoride brought about the characteristic abnormalities" in rats. However, McClure and Mitchell (69), and Mitchell (77), observed no outward changes in the teeth of swine as a result of the presence of fluorides in the ration.

Effect on other constituents. Experiments have been reported which indicate that a change in structure and in the content of elements other than F may occur in the bones of animals fed fluorides. McClure and Mitchell (68) found "that sodium fluoride at levels of approximately 0.03 per cent and 0.06 per cent fluorine caused a consistent increase,

averaging 1.30 per cent, in the ash content of the bones" and "a statistically significant, though not as consistent a decrease, in the calcium content of the ash, averaging 1.05 per cent. The phosphorus content of the ash was not significantly affected, but the ratio of calcium to phosphorus in the ash was depressed." There was a tendency for CaF_2 at the same levels of F to increase the ash in the bones and to depress the Ca in the ash of the femurs but not in the ash of the humeri.

Contrary to the above findings as applied to swine nutrition, Bethke and co-workers (37) found no changes in the total ash, or in the Ca and P content of the ash following additions of fluorides to the ration of young growing pigs.

Schwyzer (62) compared the total Cl content of fresh bone of normal dogs with the Cl in bones from two dogs fed fluorides. His results do not include a sufficient number of individuals to warrant any conclusions regarding the effect of F. Likewise the report of the Ca content and specific gravity (61) of the bones is incomplete and cannot be interpreted as indicating any differences due to F.

Mitchell (77) studied the composition of the humeri and scapulas of pigs as affected by CaF_2 and by rock phosphate additions to the ration. As regards the effects of feeding CaF_2 in these experiments, the most that may be said is that there is rather strong evidence of a decrease in the Ca and P in the ash of the humeri. This is consistent with the finding of McClure and Mitchell as regards the Ca and P of the ash of bones of rats fed fluorides. Bethke, it is recalled, observed no such changes in the bones of pigs. Mitchell's results do not warrant conclusions as regards Ca in the water-free and fat-free bones.

Mottled enamel. A defect in human teeth known as mottled enamel is endemic in certain states in the United States, in South America, China, Holland, Spain, Italy, Mexico, South Africa, Bahama and the Cape Verde Islands (78). The cause of this abnormality has been attributed to the water supply, particularly by McCay (78). Recently Smith, Lantz and Smith (20) presented the first experimental evidence that F contained in drinking water is probably the causal factor. Churchill (21) has found F present in waters from localities affected with mottled enamel. Smith, Lantz and Smith investigated endemic mottled enamel occurring in St. David, Arizona. Examination was made of the teeth of children in elementary grades and high school, and subsequently the occupants of nearly every home in the community were examined. According to these investigators "every child exposed to the environmental conditions of this community during the years of growth

of the enamel of the permanent teeth is certain to have mottled enamel." The fact that the water supply is the source of the trouble appears to have been proved by the production of a similar condition in the enamel of experimental rats by giving the rats St. David water to drink or by giving food in which the residue from St. David water had been incorporated. "The mottling of the enamel thus produced has been found to be identical with that resulting from the inclusion of sodium fluoride in the ration." Quantitative analysis of St. David drinking waters revealed the presence of 3.8 to 7.15 mgm. of F per liter as compared with 0 to 0.3 mgm. in waters from non-endemic communities.

Mottled enamel in human teeth is characterized by chalky white patches distributed over the surface; frequently the enamel is pitted and corroded. There is an absence of normal translucency, the teeth showing a dead white unglazed appearance. Histological examinations according to Black (79) and Leon (80) indicate imperfect calcification particularly of the enamel rods themselves. The intercementing material which is normally present between the enamel rods is lacking.

It may be concluded that fluorides produce pronounced and characteristic structural changes in teeth. There is evidence that bones may also be affected structurally by fluorides. Apparently F is present in excessive amounts in the bones and teeth of animals fed fluorides. There is evidence that the ash, Ca and P content of bones may be affected detrimentally by the ingestion of small amounts of fluorides. From the results of their experiments with rats, McClure and Mitchell concluded "that fluorine, especially when consumed in the more soluble form of the sodium salt, may cause the deposition of an apparently abnormal constituent in the bones, or an abnormal deposition of a non-calcium constituent, possibly a fluoride of a mineral other than calcium, as evidenced by an increase in percentage of ash in the bones above the normal. There is also an interference with the deposition of calcium in the bones brought about by feeding fluorine, a result in agreement with the calcium balance studies previously described." At what appear to be extremely low levels of F intake, these defects may be manifested in the growth and outward appearance of the teeth.

c. *Effect of ingestion of small amounts of fluorine on food consumption, growth, and reproduction.* Chronic F intoxication in albino rats as studied by Sollman, Schettler and Wetzel (63), included the feeding of NaF, NaCl, "phosphate rock," $Ca_3(PO_4)_2$ (precipitated) and $CaCO_3$ (precipitated). Doses of pure NaF, ranging from 0.0015 mgm. to 151.0 mgm. per kgm. of body weight daily, were added to an otherwise com-

plete ration, and the effect on food consumption and growth was observed. Generally, the results obtained show a parallelism between food intake and growth. In three instances, however, the food consumption was normal or above, whereas the rate of growth was below normal. The amount of F in the rations in these cases varied from 0.018 to 0.0452 per cent. These results may be regarded as indicating an effect of F at these levels of intake on the rate of growth of rats. An amount of F equivalent to 0.103 per cent of the ration proved fatal to seven rats within 11 weeks. Six rats received a ration containing 0.045 per cent F, along with a normal ration, both fed *ad libitum*. Three of these rats died by the end of the fifth week. Since the results show that all the fatalities occurred on the higher level of F feeding, it may be concluded that F was the causal factor.

Schulz and Lamb (66) added NaF to a ration which had proved adequate for normal growth and reproduction of rats, in amounts representing 0.0226, 0.0452, 0.0678, and 0.1130 per cent F. It is evident that the higher amount of F is definitely toxic, but it is difficult to interpret the effects of the lower amounts in the absence of data concerning the food intake and the rate of growth of the individual rats. Regarding the effect on reproduction the authors state, "It appears that an unfavorable effect on reproduction begins at a level of about 0.0250 per cent NaF (0.0102 per cent F) in the ration." Confirming this observation as regards reproduction in rats, Del Castillo (81) and Chaneles (70) report that fluorides interfere with the estrual cycle of rats.

Rats fed daily doses of 0.002–0.003 gram of NaF by Goldemberg (49) showed greatly retarded growth. This investigator chooses to call the effect a *cretinisme fluorique*, even suggesting that endemic cretinism may be due to F contained in the soil, air, food and water, of affected regions. Goldemberg's work would appear to lack proper control and his assumptions need further experimental confirmation, particularly as regards endemic cretinism and also endemic goiter. (His observations with regard to the effect of F on the thyroid gland will be mentioned later.)

According to investigators at the Iowa Agricultural Experiment Station (82) (83), the results of studies on the effect of fluorides in the ration upon the growth and reproduction of albino rats justify the following conclusions: "the toxic effects produced by sodium fluoride vary directly with the amount of calcium salts in the ration;" "the toxic effects of calcium fluoride are much less marked than those of sodium fluoride." "Normal growth and reproduction proceeded even when the

concentration of CaF_2 was increased to 0.25 per cent (0.113 per cent F) in a ration composed of corn, tankage, linseed oil meal, bone meal, and salt. In more highly purified rations which contained wheat, limestone, and butter fat, growth and reproduction failed when the concentration of CaF_2 was at a lower level."

McClure and Mitchell (68) studied primarily the effect of F on growth, bone and tooth development and the utilization of Ca by the experimental animals. Young growing rats were fed NaF by the paired feeding method, in an otherwise normal ration, in amounts equivalent to approximately 0.01, 0.03 and 0.06 per cent F. Two levels of CaF_2 providing about 0.03 per cent and 0.06 per cent F were also fed to growing rats. Inspection of the growth curves as well as a statistical analysis of the growth data warrants the conclusion "That calcium fluoride as well as sodium fluoride at levels of 0.0313 per cent and 0.0623 per cent fluorine in the ration inhibit growth, entirely aside from any effect on food consumption." "The higher level of sodium fluoride appears to have lowered the food consumption of the rats also, but the corresponding level of calcium fluoride had no apparent influence on food consumption."

In dairy rations. In an attempt to duplicate toxic effects obtained by feeding raw rock phosphate to dairy cattle, Taylor (67) at the Michigan Agricultural Experiment Station has fed several dairy cows varying amounts of F in the form of calcium fluosilicate, $CaSiF_6 \cdot 2H_2O$. One animal fed F at a level of 0.125 per cent of the dry matter of the ration, went off feed at once and "refused to eat either its roughage or grain mixture in any appreciable amounts for a period of three weeks." At this time the amount of F was reduced to 0.4 gram daily, and the ration was properly consumed. Five grams of $CaSiF_6 \cdot 2H_2O$ (2.28 grams of F) fed daily to one cow caused no deleterious effects until after the cow freshened, when there was a rapid decline in weight and a loss of appetite. The above response to F feeding has been repeated in other dairy cows by the same investigator, and the conclusion "that fluorine is detrimental to health and general condition of dairy animals" at the levels fed seems justified.

In swine rations. The experiments of Bethke, Kick, Edgington and Wilder (37) show that F in the form of NaF, when added to the ration of growing swine, among other effects lowers the food consumption. Since the data show that the average daily gain among the different lots of pigs varies directly with the food consumption, there is difficulty in assessing the rôle of F in the growth response observed in these experiments.

Concurrent with their above mentioned experiments with rats, McClure and Mitchell studied the effect of CaF_2 and phosphate rock on the calcium retention of young growing pigs (69). The results of feeding fluorides to three pigs at levels of about 0.013, 0.017, and 0.026 per cent F (0.0032, 0.0035, and 0.0055 gram F per kgm. of body weight respectively), are regarded as having demonstrated a detrimental effect upon the food consumption and growth of these pigs.

There is unmistakable evidence that food consumption is less and the rate of growth is retarded, due to the ingestion of small quantities of fluorides. The evidence indicates that levels of F as low as 0.03 per cent F in the ration may inhibit growth of rats. It is possible that the form of F as well as other constituents of the ration affect the toxic level. Apparently the dairy cow, as indicated by the general nutritive state as well as by the refusal to consume food, is particularly susceptible to the effects of very small amounts of F. The evidence is less striking with regard to the response of swine. However, the experiments with rock phosphate in which swine have been used as the experimental animals, indicate toxic disturbances, which in the present state of our knowledge, are generally attributed to F contained in the rock phosphate. There is evidence to indicate that 0.01 per cent F in the ration may interfere with normal reproduction in rats.

d. *Effect of fluorine on enzymatic changes: Reactions of lipase.* An inhibitory effect on enzymes characteristic of antiseptics in general is a property of ordinary inorganic fluorides. Experimental evidence has established the fact that there is also a specific influence of fluorides on certain enzymatic changes associated particularly with carbohydrates and fats. Thus the results of a systematic study conducted by Kastle and Loevenhart (84) on the effect of antiseptics on the reactions of pancreatic and liver extracts revealed a harmful effect of most substances studied and also a particularly remarkable destructive action of NaF on the reactions of lipase. Subsequently Loevenhart and Peirce (85) investigated, among other substances, the halides of Na and K; the chlorides of Ca, Cd, Ba and Mn; $NaNO_3$ and KNO_3; Na_2HPO_4; K_2CrO_4; and NH_4CNS, but none of these substances showed an inhibitory effect comparable to the action of F. Dilutions of NaF as low as 1:15,000,000 may inhibit the action of lipase on ethyl acetate as much as 50 per cent. With ascending order of the fatty acid series the inhibition continues to decrease. Leake et al. (86) have obtained evidence that NaF also inhibits the action of this enzyme *in vivo*.

As regards the mechanism of this fluoride effect Amberg and Loevenhart (87) state that the facts indicate "that the fluoride probably acts by

retarding the formation or decomposition of some intermediate product formed in the hydrolysis of the ester by the enzyme." That the action of NaF may be explained as a monomolecular combination of the enzyme with part of the NaF is the conclusion of Peirce (88) and of Murray (89). Rothschild (90) on the other hand believes that F by no means reacts with the free enzyme, though this does not mean that the effect is non-specific. He suggests the possibility of a combination of F with the enzyme-substrate complex.

On carbohydrate enzymes: Certain phases of the enzymatic breakdown of carbohydrate are particularly sensitive to an effect of the fluoride ion. The activity of the amylase of the potato as measured by its ability to produce invert sugar is increased by solutions of NaF up to 1/1.75 molar concentration (91). A study of the effect of oxalate (92), and fluoride (121), on glycolysis in blood, has revealed a strongly inhibitory effect. Lang and Lang (93) found that pancreatic diastase, as contained in an extract of beef pancreas, formed less maltose and a relatively greater amount of glucose in the presence of fluoride. It appeared to these authors that, while the total diastatic effect was inhibited by fluoride due to a decrease in breakdown of starch to maltose, there was a stimulating effect of fluoride on the maltose to glucose change. The work of Clifford (94) indicates an effect of fluorides on the enzymatic decomposition of starch. Clifford determined the time required for the achromatic point to appear in a 0.5 per cent solution of starch treated with 2.0 cc. of saliva and various salt solutions. She reports that the Cl ion accelerates the reaction, the Br ion is inert, while the reactions of the fluorides and iodides are specific effects of the salts studied, i.e., NaF is inert while KF and NH_4F show a marked slowing effect.

Glycolysis in muscle: Information as to the action of fluorides on glycolysis has been extended largely by the research conducted by Embden and his school (95) (96) (97) (98), concerning the influence exerted by ions upon the chemical changes associated with muscle metabolism. It became evident after certain experiments that tartrate and fluoride inhibit or even reverse the glycogen breakdown changes of muscle. Of these two ions, fluoride is most potent and has received the greatest attention from the standpoint of the nature of the fluoride effect as well as for the elucidation of the nature of the chemical changes involved in muscle metabolism.

The early work of Embden and co-workers drew attention to the fact that F decreased the amount of free phosphoric acid formed in muscle

pulps suspended in buffer solutions. It was supposed that the formation of a phosphate ester, described as lactacidogen, caused this decrease. The evidence indicated that the action of F might be a synthetic one, as for example Embden and Haymann found, with the disappearance of the free phosphoric acid, a definite formation of diphosphoric ester. This synthetic action can be increased by addition of glycogen to the fluoride solution (99) (100). Larger quantities of added phosphate may be synthesized if sufficient glycogen is added (101) (102).

Recent research, generally, has not confirmed Embden's original idea of the nature of the synthetic effect of fluoride. Thus Eggleton and Eggleton (103), having studied the effect of the fluoride ion as discovered by Embden and Lehnartz, state, "The 'synthesis' observed by these authors was in the main a conversion of phosphagen-phosphorus (which they estimated as inorganic P) into some acid-stable phosphoric ester. Where phosphagen is present some true synthesis can occur, however, for added inorganic phosphate disappears to a small extent under the action of fluoride." (Eggleton (104) has recently presented a review of the rôle of phosphorus in the chemical mechanism of muscular contraction, which includes a discussion of the developments of this phase of the subject.)

It is known that fluorides do not prevent a decrease in glycogen of muscle pulp undergoing fermentative changes but do prevent the formation of lactic acid (105) (106) (107). Fluorides also inhibit alcoholic fermentation by yeast (108), an observation which is not surprising, in view of the many analogies between lactic acid fermentation in muscle extract and alcoholic fermentation (a notable example of this similarity is the occurrence of practically the same hexosephosphates). Lipman (108) sought to determine whether this inhibitory effect of F might not be due to a retarding of the hydrolysis of hexose-phosphorus-acid-ester, a compound which is formed in the enzymatic decomposition of glycogen. Meyerhof (109), as well as Lipman (110), obtained evidence that the accumulation of hexose-diphosphoric acid, as observed by Embden under the influence of fluoride, is due to the inhibition of the dissociation of the intermediary-formed phosphoric-acid-ester, which in the absence of lactic acid formation goes over into a more enzyme-resistant form, i.e., hexose-diphosphoric acid. The strong inhibition is due to the position of F in the lyotropic series. The decrease in lactic acid formation is due to the inhibition of the hexose-phosphate breakdown. Lipman (110), after studying carbohydrate fermentation by yeast as inhibited by fluorides, concluded that the inhibition is

reversible, that the extent of the inhibition for different fluoride concentrations is in agreement with the law of mass action, and that the inhibition is greater with increased acidity. These properties are similar to those observed in connection with fluoride inhibition of lipase and agree with the assumption that F forms an easily dissociated compound with the enzyme, thus preventing the normal dissociation of hexose-phosphate. Confirming this explanation is the fact shown by Lipman (110), that the hydrolysis of glycerophosphoric acid, hexose diphosphoric acid and hexose monophosphoric acid is inhibited only slightly less by F than is lactic acid formation.

Lipman (110) presents some evidence which indicates the formation of a complex compound of iron with F. Support for this view comes from a comparison of the above reactions with the dissociation curve of fluoro-methemoglobin.

Dickens and Simer (111) have confirmed the essential facts brought out by Lipman and conclude in part "that fluoride forms an inactive compound with some substance essential for glycolysis. This reaction follows the ordinary laws of equilibria." Lohmann (112), in a recent contribution to the subject, states that the effect of fluorides, oxalates and citrates in lactic acid forming reactions of muscles, consists of the inhibition of the splitting off of phosphoric acid. As a result of this condition there is a change of first-formed (primary) hexose phosphoric acid ester into a difficultly hydrolysable ester.

Pepsin and urease: Few striking effects of F on proteolytic enzymes have been demonstrated. Thus the digestion of white of egg, blood fibrin and protein of cow's milk by pepsin and trypsin was found uninfluenced by the presence of fluorides (113). Von Groer (114) obtained some evidence that the gelatinase of Bacillus prodigiosus was retarded in its activity by the presence of a 1 per cent solution of NaF. There is unmistakable evidence, however, presented by Clifford (115) (116), that the clotting time of milk under the influence of pepsin is retarded by fluorides. Inhibition may also be brought about by iodides but the reaction is hastened by chlorides and bromides. The halides of Na, K and NH_4 were studied by Clifford.

Jacoby (117) first observed that F inhibits the reaction of urease. Subsequent research (118) (119) by this investigator has brought out the following facts regarding the urease effect: *a*, F is most effective at optimum conditions for enzyme reactions; *b*, it is most active against urease in acid solution; *c*, the effect is more pronounced in buffered solutions on account of the reaction being influenced by the ammonia

formed in the hydrolysis; *d*, the F effect cannot be duplicated with oxalic and citric acids; *e*, iodine salts are also very active against urease.

e. *Additional body tissues and functions affected by fluorine.* Attempts to demonstrate an effect of F on the composition of various tissues (not including bones and teeth), other than an increase in the F content, have been generally unsuccessful. Trebitsch (36), as well as Brandl and Tappeiner (29), has reported the deposition of abnormal amounts of F in various body tissues by continuous small additions of F to the ration. Cristiani and Gautier (50) fed varying amounts (0.1–1.0 gram daily) of NaF to a dog for a period of about 21½ months, and there resulted what appear to be abnormal amounts of F deposited in various tissues. Schwyzer (62) reports an effect of F on some of the form elements of blood. Whereas 55–60 per cent of the leucocytes of the blood of a normal rabbit are multinuclear, in the blood of rabbits fed F these values range from 10–34 per cent. He believes these changes in the leucocytes are due to an irritation or inflammation of the bone marrow. He noted a partial change of the yellow bone marrow into red.

On calcium metabolism: Gerschmann (120) reports that the injection of 30 mgm. NaF per kgm. body weight reduces blood Ca in 24 hours. It cannot be said that Schwyzer (62) was successful in showing an effect of F on Ca metabolism. Direct experimental evidence of an effect of ingested F compounds on Ca metabolism was obtained by McClure and Mitchell (68) (69). These investigators saw fit to conclude that "the calcium balance of the rats was not affected by either sodium fluoride or calcium fluoride at levels of 0.0106 per cent and 0.0313 per cent F in the ration. The results indicate with a probability equal to practical certainty, that a level of 0.0623 per cent of fluorine in the form of either calcium fluoride or sodium fluoride, lowered the percentage of calcium that otherwise would have been retained by the experimental rats" (68). The results of analyses show definitely a depression of the Ca content of the ash of bones of rats fed NaF and therefore may be regarded as further evidence of an effect of F on Ca metabolism.

Clotting of blood: It has been established that there is a small amount of F present in normal blood and that ingested fluorides may increase significantly this amount. Stuber and Lang (121) observed a number of cases of hemophilic patients with amounts of blood F so abnormally high, that they were led to suspect F as being a causal factor. In two patients they found clotting times of 150 and 180 minutes associated with 0.0029 per cent and 0.0039 per cent F, respectively, present in the blood. It was recalled by these investigators that normal goose and

rabbit blood clots slowly, and according to Stuber and Lang (121) relatively large amounts of F occur in the blood of these two species. Cat blood and dog blood clots rapidly and was found free of F by Stuber and Lang.

Continuing their observations relative to the blood clotting process, these investigators discovered that the blood of normal individuals residing in the town of Freiburg, Germany, would clot in from six to nine minutes, whereas in the region of Kiel, the clotting time varied from ten to seventeen minutes. Their results of analysis for F in the blood of Kiel "patients" show much variation. In the absence of a report of the F content of blood from Freiburg residents it is impossible to assess the significance of their data. Apparently in numerous instances a large amount of F has accompanied a long clotting period. Analysis of Kiel tapwater revealed a relatively high content of fluorine as compared with Freiburg water, which contained no fluorine. Lang (122), however, found that ingested organic fluorine compounds had no effect on blood fluorine or coagulation time.

According to Stuber and Lang the rate of glycolysis in the blood of Kiel patients was less, due to the presence of fluoride. These investigators are of the opinion that a retarding of glycolysis may affect indirectly the clotting power.

Contrary to the results of Stuber and Lang, Schwyzer (62) observed a very high degree of coagulability of blood of rabbits, pigeons and dogs which were fed F. Whereas the coagulation of the total blood of a normal rabbit in an open vessel required from 6 to 8 minutes, the coagulation time in the case of fluorine-rabbits was 4 minutes in one instance and $2\frac{1}{2}$ minutes in another. Blood in a capillary coagulated in 5 minutes for a normal rabbit and in 50 seconds, $1\frac{1}{2}$ minutes, 3 minutes, $3\frac{1}{2}$ minutes, and $2\frac{3}{4}$ minutes in the case of six rabbits fed fluoride.

On muscular irritability. A form of toxicity of F directly ascribable to its ability to precipitate calcium is described by J. Loeb (123). Loeb put a fresh muscle (gastrocnemius) of a frog into dilute solutions of NaF, Na_2CO_3,Na_2HPO_4, sodium oxalate, sodium citrate and sodium tartrate; under this condition the muscle as a rule shows no reaction. When taken out of the salt solutions into the air, CO_2, oil, sugar solutions, etc., it begins to contract powerfully. If the nerve alone (without the muscle) is put into one of these salt solutions "the muscle begins to twitch in about five minutes and finally goes into tetanus." While Loeb believes that the diminution of Ca-ions is the cause of the effect of NaF and similar solutions, he considers it possible "that these

solutions may have other effects which play a rôle in these phenomena." This form of F irritability has not received further investigation except as it is related to a disturbance of the normal ratio of salts in the muscle (124) (125).

Effect on thyroid gland: Several reports appear in the literature to the effect that fluorine causes an enlargement of the thyroid gland. Maumeme (126) observed a swelling in the neck of a dog fed fluoride. His findings were not confirmed by other investigators.

Goldemberg (49) goes so far as to suggest that excess fluorine in the diet may be a cause of endemic goiter. He claims that by the ingestion or injection of 2 to 3 mgm. of fluorine per day for six months the thyroid gland in white rats may be increased 5 to 6 times in volume. Cristiani (127) studied the condition of the thyroid glands of animals which had died of chronic fluorine poisoning. In the case of fourteen such animals (guinea pigs) more or less marked lesions were observed in the thyroid glands. The alterations consisted essentially of a proliferation of the parenchymatous tissue, less rarely of the interstitial tissue.

BIBLIOGRAPHY

(1) Steiger, G. Journ. Amer. Chem. Soc., 1908, xxx, 219.
(2) Merwin, H. E. Amer. Journ. Sci., series 4, 1909, xxviii, 119.
(3) Gautier, A. and P. Clausmann. Compt. rend. Acad. Sci., 1912, cliv, 1469.
(4) Gautier, A. and P. Clausmann. Compt. rend. Acad. Sci., 1912, cliv, 1753.
(5) Gautier, A. and P. Clausmann. Compt. rend. Acad. Sci., 1912, cliv, 1670.
(6) Berzelius, J. Journ. f. Chemie u. Physik, 1816, xvi, 423.
(7) Rose, H. Annalen der Chemie und Pharmacie von F. Wohler und J. Liebig, 1849, lxxii, 343.
(8) Gooch, F. A. and M. Kobayashi. Amer. Journ. Sci., series 4, 1918, xlv, 370.
(9) Meyer, R. J. and W. Schulz. Zeitschr. f. angew. Chem., 1925, xxxviii, 203.
(10) Fresenius, R. Zeitschr. f. analyt. Chem., 1866, v, 190.
(11) Tammann, G. Zeitschr. f. analyt. Chem., 1885, xxiv, 328.
(12) Offermann, H. Zeitschr. angew. Chem., 1890, iii, 615.
(13) Wagner, C. R. and W. H. Ross. Journ. Ind. Eng. Chem., 1917, ix, 1116.
(14) Reynolds, D. S., W. H. Ross and K. D. Jacob. Journ. Assoc. Off. Agri. Chem., 1928, xi, 225.
(15) Steinkoenig, L. A. Journ. Ind. and Eng. Chem., 1919, xi, 463.
(16) Jacob, K. D. and D. S. Reynolds. Journ. Assoc. Off. Agri. Chem., 1928, xi, 237.
(17) Carles, P. Compt. rend. Acad. Sci., 1907, cxliv, 37.
(18) Gautier, A. and P. Clausmann. Compt. rend. Acad. Sci., 1914, clviii, 1389.
(19) Carnot, A. Compt. rend. Acad. Sci., 1896, cxxiii, 724.
(20) Smith, M. C., E. M. Lantz and H. V. Smith. Arizona Agri. Exp. Sta. Technical Bulletin, no. 32, June 10, 1931.

(21) Churchill, H. V. Ind. and Eng. Chem., 1931, xxiii, 996.
(22) Gautier, A. and P. Clausmann. Compt. rend. Acad. Sci., 1916, clxii, 105.
(23) Woodman, A. G. and H. P. Talbot. Journ. Amer. Chem. Soc., 1907, xxix, 1362.
(24) Gay Lussac, De M. and Berthollet. Ann. de chimie, 1805, lv (series 1), 258.
(25) Berzelius. Ann. de chimie, 1807, lxi, 256.
(26) Weiske, H. Landwirtschaft Versuch. Stat., 1889, xxxvi, 81.
(27) Carnot, A. Compt. rend. Acad. Sci. 1892, cxiv, 1189.
(28) Zalesky. Med-chem. Untersuchungen von Hoppe-Seyler, 1866–71, i, 19.
(29) Brandl, J. and H. Tappeiner. Zeitschr. f. Biol., 1891, xxviii, 518.
(30) Gabriel, S. Zeitschr. f. Physiol. Chem., 1894, xviii, 257.
(31) Wrampelmeyer, E. Zeitschr. f. Analyt. Chem., 1893, xxxii, 550.
(32) Jodlbauer. Zeitschr. f. Biol., 1903, xliv, 259.
(33) Gassman, T. Zeitschr. f. Physiol. Chem., 1908, lv, 455.
(34) Gautier, A. and P. Clausmann. Compt. rend. Acad. Sci., 1913, clvi, 1425.
(35) Gautier, A. and P. Clausmann. Compt. rend. Acad. Sci., 1913, clvi, 1347.
(36) Trebitsch, F. Biochem. Zeitschr., 1927, cxci, 234.
(37) Bethke, R. M., C. H. Kick, B. H. Edgington and O. H. Wilder. Amer. Soc. Anim. Prod., Proc. 1929, p. 29.
(38) Gautier, A. and P. Clausmann. Compt. rend. Acad. Sci., 1913, clvii, 94.
(39) Wilson, G. Chem. Gax., 1850, p. 365. In Forschritte der reinen und tech. Chemie von Just. Liebig und H. Kopp. 1850, iii, 278.
(40) Tamman, G. Zeitschr. f. Physiol. Chem., 1888, xii, 322.
(41) Stuber, B. and K. Lang. Biochem. Zeitschr., 1929, ccxii, 96.
(42) Schulz, H. Arch. f. exp. Path. u. Pharm., 1889, xxv, 326.
(43) Tappeiner, H. Arch. f. exp. Path. u. Pharm., 1889, xxv, 203.
(44) Gautier, A. Compt. rend. Acad. Sci., 1914, clviii, 159.
(45) Shear, M. J. and B. Kramer. Journ. Biol. Chem., 1928, lxxix, 125.
(46) Baldwin, H. B. Journ. Amer. Chem. Soc., 1899, xxi, 517.
(47) McNallay, W. D. Journ. Amer. Med. Assoc., 1923, lxxxi, 811.
(48) Goldemberg, L. Journ. de Physiol. et de Path. Gen., 1930, xxviii, 556.
(49) Goldemberg. L. Journ. de Physiol. et de Path. Gen., 1927, xxv, 65.
(50) Cristiani, H. and R. Gautier. Compt. rend. de la Soc. de Biol., 1925, xcii, 139.
(51) Cristiani, H. and R. Gautier. Compt. rend. de la. Soc. de Biol., 1925, xcii, 946.
(52) Cristiani, H. and R. Gautier. Compt. rend. de la Soc. de Biol., 1925, xcii, 1276.
(53) Cristiani, H. and A. Gautier. Compt. rend. de la. Soc. de Biol., 1925, xciii, 911.
(54) Cristiani, H. and P. Chausse. Compt. rend. de la Soc. de Biol., 1927, xcvi, 842.
(55) Cristiani, H. and P. Chausse. Compt. rend. de la Soc. de Biol., 1927, xcvi, 843.
(56) Cristiani, H. Compt. rend. de la Soc. de Biol., 1927, xcvi, 388.
(57) Cristiani, H. Compt. rend. de la. Soc. de Biol., 1930, ciii, 292.
(58) Sette, N. Compt. rend. de la Soc. de Biol., 1928, xcviii, 1094.

(59) Sonntag, G. Arb. kais. Gsndhtsamte. Amst., 1916, l, 307.
Chem. Zentr., 1916, lxxxvii (87), 1095.
(61) Schwyzer, F. Journ. Med. Res., 1903, x, 301.
(62) Schwyzer, F. Biochem. Zeitschr., 1913, lx, 32.
(63) Sollmann, T., O. H. Schettler, N. C. Wetzel. Journ. Pharm. Exper. Therap. 1921, xvii, 197.
(64) Marcovitch, S. Univ. Tennessee Agric. Exp. Sta., Bull. no. 139, 1928, 48 pages.
(65) McCollum, E. V., N. Simmonds, J. E. Becker and R. W. Bunting. Journ. Biol. Chem., 1925, lxiii, 553.
(66) Schultz, J. A. and A. R. Lamb. Science, 1925, lxi, 93.
(67) Taylor, G. E. The Quart. Bull., Michigan Agric. Exp. Sta., 1925, xi, 101.
(68) McClure, F. J. and H. H. Mitchell. Journ. Biol. Chem., 1931, xc, 297.
(69) McClure, F. J. and H. H. Mitchell. Journ. Agric. Res., 1931, xlii, 363.
(70) Chaneles, J. Compt. rend de la Soc. Biol. 1929, cii, 860.
(71) Lewy, A. Arch. Otolaryngol., 1928, viii, 315.
(72) Tolle, C. and L. H. Maynard. Record of Proceedings, Annual Meeting Amer. Soc. Animal Production 1928, p. 15.
(73) Arkansas Agricultural Exp. Station 38th Annual Report 1926, Bull. no. 215, p. 23.
(74) Bergara, C. Compt. rend. Soc. Biol., 1927, xcvii, 600.
(75) Reed, O. E. and C. F. Huffman. Mich. Agric. Exp. Sta., Quart. Bull., 1928, x, 152.
(76) Reed, O. E. and C. F. Huffman. Mich. Agric. Exp. Sta., Tech. Bull., 1930, cv, 63 p.
(77) Mitchell, H. H. Unpublished data from the University of Illinois.
(78) McKay, F. S. Journ. Dental Res., 1930, x, 561.
(79) Black, G. V. Dental Cosmos, 1916, lviii, 129.
(80) Leon, W. J. Journ. Dental Res., 1918, v, 117.
(81) Del Castillo, E. B. Compt. rend. Soc. Biol., 1928, xcix, 1405.
(82) Iowa Agric. Exp. Station Report for year ending June 30, 1928, p. 29.
(83) Iowa Agric. Exp. Station Report for year ending June 30, 1929, p. 29.
(84) Kastle, J. H. and A. S. Loevenhart. Amer. Chem. Journ., 1900, xxiv, 491.
(85) Loevenhart, A. L. and G. Peirce. Journ. Biol. Chem., 1907, ii, 397.
(86) Leake, C. D., A. D. Dulmes, D. N. Treweek and A. S. Loevenhart. Amer. Journ. Physiol., 1929, xc, 426.
(87) Amberg, S. and A. S. Loevenhart. Journ. Biol. Chem., 1908, iv, 149.
(88) Peirce, G. Journ. Biol. Chem., 1913–14, xvi, 5.
(89) Murray, D. R. P. Biochem. Journ., xxiii, 1929, 292.
(90) Rothschild, P. Biochem. Zeitschr., 1929, ccvi, 186.
(91) Doby, G. Biochem. Zeitschr., 1914, lxvii, 166.
(92) Macleod, J. J. R. Journ. Biol. Chem., 1913, xv, 497.
(93) Lang, S. and H. Lang. Biochem. Zeitschr., 1921, cxiv, 165.
(94) Clifford, W. M. Biochem. Journ. 1925, xix, 218.
(95) Embden, G. and E. Lehnartz. Zeitschr. f. Physiol. Chem., 1924, cxxxiv, 243.
(96) Embden, G. and C. Haymann. Zeitschr. f. Physiol. Chem., 1924, cxxxvii, 154.

(97) Abraham A. and P. Kahn. Zeitschr. f. Physiol. Chem., 1924, cxli, 161.
(98) Lang, H. and M. E. Mayer. Zeitschr. f. Physiol. Chem., 1924, cxli, 181.
(99) Andrews, S. Biochem. Journ., 1925, xix, 242.
(100) Embden, G. and M. Zimmerman. Zeitschr. f. Physiol. Chem., 1924, cxli, 225.
(101) Beatie, F. and T. H. Milroy. Journ. Physiol., 1925, lx, 379.
(102) Beatie, F. and T. H. Milroy. Journ. Physiol., 1926–27, lxii, 174.
(103) Eggleton, P. and G. P. Eggleton. Journ. Physiol., 1928, lxv, 15.
(104) Eggleton, P. Physiol. Reviews, 1929, ix, 432.
(105) Lang, H. Zeitschr. f. Physiol. Chem., 1924, cxxxvii, 105.
(106) Davenport, H. A. and M. Catonio. Journ. Biol. Chem., 1927, lxxiii, 463.
(107) Euler, H. and H. Cramer. Biochem. Zeitschr., 1914, lx, 25.
(108) Lipmann, F. Biochem. Zeitschr., 1928, cxcvi, 3.
(109) Meyerhof, O. Biochem. Zeitschr., 1926, clxxviii, 462.
(110) Lipmann, F. Biochem. Zeitschr., 1928, ccvi, 171.
(111) Dickens, F. and F. Simer. Biochem. Journ., 1929, xxiii, 936.
(112) Lohmann, K. Biochem. Zeitschr., 1930, ccxxii, 324
(113) Vandevelde, A. J. J. and E. Poppe. Biochem. Zeitschr., 1910, xxviii, 134.
(114) Von Groer, F. Biochem. Zeitschr., 1912, xxxviii, 252.
(115) Clifford, W. M. Biochem. Journ., 1927, xxi, 544.
(116) Clifford, W. M. Biochem. Journ., 1928, xxii, 1128.
(117) Jacoby, M. Biochem. Zeitschr., 1916, lxxiv, 107.
(118) Jacoby, M. Biochem. Zeitschr. 1928, cxcviii, 163.
(119) Jacoby, M. Biochem. Zeitschr., 1929, ccxiv, 368.
(120) Gerschmann, R. Ann. farm. bioquim., 1930, i, 77, in C. A. 1931, xxv, 348.
(121) Stuber, B. and K. Lang. Biochem. Zeitschr., 1929, ccxii, 96.
(122) Lang, H. Arch. exp. Path. u. Pharm. 1930, clii, 361.
(123) Loeb, J. Amer. Journ. Physiol., 1901, v, 362.
(124) Zoethout, W. D. Amer. Journ. Physiol., 1902, vii, 320.
(125) Garrey, W. E. Amer. Journ. Physiol., 1905, xiii, 186.
(126) Maumeme. Compt. rend. Acad. Sci., 1866, lxii, 381.
(127) Cristiani, H. Compt. rend. Soc. Biol., 1930, ciii, 554.

THE STRUCTURAL CHANGES IN STRIPED MUSCLE DURING CONTRACTION

H. E. JORDAN

School of Histology and Embryology, University of Virginia

The detailed structure of striped muscle in full contraction differs markedly from that in the relaxed condition. During the intervening functional phases the structure of the component fibers changes in orderly progressive stages from one extreme to the other. Functional states are strictly correlated with definite structural features. Accordingly, a primary consideration of an adequate explanation of muscle contraction is that it must be consistent with alterations in the relationships of the constituent structural elements of the fiber. An understanding of the structural changes during contraction requires a preliminary description of the minute composition of the fiber at rest.

Microscopic anatomy. Skeletal muscle consists of fibers held together by connective tissue. This continuous connective tissue comprises the enveloping epimysium; that portion surrounding the constituent bundles of fibers, the perimysium; and that continued within the bundles and more or less closely associated with the individual fibers, the endomysium.

Development. Each fiber is enclosed by a relatively robust and resistant lamina, the sarcolemma, representing a cell membrane. The muscle fiber is the product of fusion of originally discrete myoblasts. The nuclei of the fiber, at first placed centrally, migrate peripherally as the fiber grows; definitively they are located immediately beneath the sarcolemma. As the fiber develops the delicate membranes of the component myoblasts coarsen to form the definitive sarcolemma. The fact that the sarcolemma contains no collagen, as demonstrated by Chittenden (12), is opposed to the assumption that the thickening of this membrane is the result of connective tissue accretions. According to Cameron (10) the sarcolemma appears late in development and is produced *in situ*, without relation to ingrowing mesenchyme. The definitive coarse sarcolemma is apparently the product of internal cytoplasmic additions to the original delicate myoblast membranes.

The definitive multinucleated fiber, in fact a polynucleated fusiform muscle giant-cell, is the result in part of the contribution by the constituent myoblasts of one nucleus each, and the subsequent mitotic and later amitotic division of these original nuclei. The fiber as thus constituted contains the myofibrillae contributed by the fusing myoblasts. These fibrils are aggregated in small bundles, each bundle representing the longitudinal fission products of an original myofibril. This mode of fibril multiplication is especially well illustrated in the case of certain teleosts, as first described by Heidenhain (32) for the striped muscle of the trout. The resulting bundles constitute the columns of Kölliker, in transverse section designated the areas of Cohnheim. The intervening non-fibrillar protoplasm composes the so-called sarcoplasm. Since the myofibrils represent differentiation products of the original myoplasm, it would seem in accord with a more precise terminology to speak of intra- and inter-fibrillar sarcoplasm; or, perhaps even better, the protoplasm of the fibrils (sarcostyles) might be designated sarcoplasm, that of the interfibrillar regions myoplasm.

Histology. The fibers are characterized by a uniform succession of transverse stripes. This cross striation or "banding" is the result of two chief factors: 1, the alternation of dark and light discs in the constituent myofibrils, and 2, the presence of a relatively robust and inelastic membrane, the membrane of Krause, which holds identical portions of the component fibrils at the same horizontal level. In the latter process the more delicate membrane of Heidenhain may possibly coöperate. Peripherally both membranes are firmly attached to the sarcolemma, and where nuclei intervene, to the nuclear membrane. The firm attachment of Krause's membrane to the sarcolemma accounts for the characteristic festooned condition of the sarcolemma in fixed tissue and in the artificially contracted fiber. The spinous contour of the nuclei is explained in the same way. The membrane of Krause, designated the Z membrane, bisects the clear discs of the fibrils, and interconnects adjacent fibrils; the membrane of Heidenhain, the M membrane, bisects the dim discs of the fibrils. In stained preparations the Z membrane may have the appearance of a row of granules. The "granules" mark the points of connection between the membrane and the fibrils.

In considering striped muscle structure it will be sufficient, accordingly, to describe the structure of one fibril between two successive Z membranes, keeping in mind, of course, the interfibrillar connecting portions of the Z and M membranes. The interfibrillar granular sarcoplasm probably has only trophic significance; the fibrils are the essential

contractile elements. The fibril is simply a succession of these Z bounded segments, the so-called sarcomeres or inokommata.

Function. The fiber as a whole is a collection of identical fibrils held at precise horizontal levels by the Z membranes; and the muscle as a whole is a collection of larger and smaller bundles, the muscle fasciculi, of essentially identical fibers, the entire complex held together by binding and enveloping connective tissue. Muscle structure is, therefore, simply a summation of identical elementary units, the sarcomeres. In like manner, muscle function may be regarded a summation of the function of these sarcomeres. Contraction is essentially a matter of the shortening and thickening of the muscle; thus, the result of the shortening and thickening of the constituent fibers; in turn, the result of a shortening and thickening of the myofibrils; in the last analysis, the result of a shortening and thickening of the sarcomeres of the fibrils. Explanation of muscle contraction also may be confined to a single sarcomere. If we can discover what actually happens in a fibril between two successive telophragmata (boundaries of a fibril sarcomere) during contraction, we will have an explanation of muscle contraction. Function is here closely correlated with structural changes. Any theory of muscle contraction which is inconsistent with the structural changes within the sarcomere must be to that extent considered inadequte.

If the foregoing is accepted, then the further description of the changes in striped muscle during contraction may be confined to the limits of a sarcomere. Since, however, these changes require a preliminary understanding of the structure at rest, the latter condition must be first described.

The resting sarcomere. The sarcomere or inokomma is bounded terminally by a telophragma or Z membrane. These terms are applicable both to the fiber and the constituent fibrils; in this discussion they are used generally as applying to the fibrils. In the case of vertebrate muscle they have reference usually to the fiber, in insect muscle to the fibrils. When striped muscle is macerated by treatment with acids or alkalies, the fiber and its fibrils break into discs bounded by the Z membranes. In other words, the maceration process effects a transverse fragmentation of the myofibrils or sarcostyles into sarcomeres. Since each sarcomere is bounded by a membrane, it follows that the telophragmata as seen in unmacerated preparations are actually double structures. Conditions are similar with respect to the M membranes, the mesophragmata. When maceration is prolonged the sarcomeres ultimately split transversely along the M membranes. Since each such ulti-

mate disc is bounded by a mesophragma on one side and a telophragma on the other, the mesophragma like the telophragma, though exceedingly more delicate, is also essentially a double membrane.

The telophragma. The recent claim of Teigs (89, 90) and D'Ancona (15) that the telophragma is in reality a spiral structure stretching from end to end of the fiber in the form of a helicoid is disproved by the simple fact that the macerated fiber splits into regular circular discs. Teigs admits that the observation of the spiral arrangement of the membranes and stripes is more readily made in well stretched muscle. This I believe explains the condition observed; it is the result of shearing stresses or oblique tensions. That striate muscle is very susceptible to mechanical distortion is illustrated in striking fashion by the tenotomy experiments of Davenport and Ranson (17). Section of the tendo Achilles caused extreme irregularities in the cross striations of the gastrocnemius muscle. In certain regions of the muscle the striations were distorted into zigzag shapes. Similar distortions occurred in myostatic contracture caused by tetanus toxin. Teigs (91) interprets the membrane of Krause as a continuation of the motor end plate, specialized for the transmission of nervous stimuli.

Metafibrils. In insect wing muscle where the constituent myofibrils are relatively coarse structures, it is possible by maceration to resolve them into minute elementary fibrils, the metafibrillae. Apparently, the longitudinal splitting process might be continued *ad infinitum.* On the basis of the evidence from insect wing muscle it seems allowable to assume that the myofibril of mammalian skeletal muscle is also resolvable into finer ultramicroscopic metafibrils.

Striations. Returning to the sarcomeres of the myofibrils it should be recalled that they consist medially of a dark disc flanked terminally by a light semi-disc (fig. 1). The light semi-disc is bounded terminally by the telophragma. The dark disc is the so-called disc of Brücke, designated the Q stripe or disc. This may show a median light disc, the median disc of Hensen, designated the H disc. The presence of this disc signifies a stage in contraction; it does not appear in the fully relaxed fibril. It may, however, be caused to appear in a stretched (extended) fibril. It is conspicuous also during early developmental stages. The definitive Q discs are formed by fusion of two originally distinct semi-discs. The Q discs, or when present the H disc, is bisected by the median membrane of Heidenhain, the M membrane or mesophragma. The Q disc appears darker both in unstained and stained preparations; in physical condition it is generally birefringent or anisotropic. The Q

disc stains readily; accordingly, it may be designated the chromophil disc in contradistinction to the chromophobe J disc.

Alternating with the dark or dim discs are the light or clear discs, the so-called intermediate discs of Krause or the J discs. These are

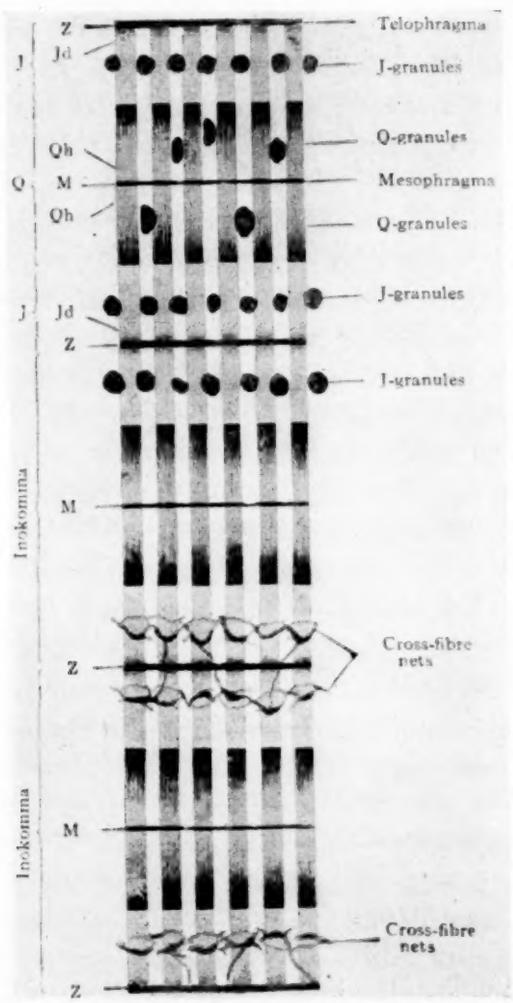

Fig. 1. Diagram of a striped muscle fiber, including six myofibrils. After Heidenhain. The terminology is referable both to the fiber and the constituent fibrils.

relatively monorefringent or isotropic. They are bisected by the telophragmata, the boundaries of the inokommata. Approximately midway between the terminal telophragma and the median Q stripe there may occur a narrow dark band. This band or disc is known as the accessory disc of Engelmann, or the N disc. Thus divided, the original J

disc includes in each half, adjacent to the telophragma, a terminal disc of Merkel, the E disc.

It follows from this description that the terms striated and striped as applied to skeletal muscle are misleading. The fundamental condition is one of alternation of physically different discs. A more precise terminology would designate skeletal muscle as "stratified" muscle, instead of banded, striped or striated.

MacCallum (60) and Menten (66) have demonstrated the aggregation of potassium, chlorides and phosphates in the Q disc. Micro-incineration likewise reveals the presence of inorganic salts in the dark discs. Scott (86) subjected both skeletal and cardiac muscle of the cat to incineration at 650°. The resulting ash was arranged in a distinctly characteristic manner, and the striations were preserved with startling clearness. The Q bands were represented by whitish ash, in some cases tinged with the yellowish red which denotes the presence of iron oxide. The J disc was ash free, except for the area of the Z membrane. In some instances the Q stripe showed a median area free of ash, indicating the H disc. The sarcolemma of the skeletal muscle was indicated by a distinct line of whitish ash in which there were some deposits of silicates.

The anisotropic band. The anisotropic condition of the Q band has been observed by many investigators. It was discovered by Goddard (28), first described in some detail by Brücke (6) and later more carefully studied by Merkel (67, 68) and Rollett (76, 77). D'Ancona (15) denies that this band is doubly refractive. While it may be confidently accepted as settled that Q is relatively anisotropic as compared with J, the fact remains that the entire sarcostyle is more or less anisotropic, varying locally in degree with the phase of contraction or relaxation. Ordinarily, however, especially in the resting condition, Q is distinctly birefringent. But anisotropy cannot in any sense be regarded as characteristic of contractile substances. Many non-contractile substances, e.g., cotton, rubber and glass, are under certain conditions anisotropic. These examples, together with the varying degree of anisotropy of the myofibril and its Q constituents, make it clear that the physical condition of birefringence cannot properly be regarded as indicative of a specific substance.

The question arises whether the Q disc appears anisotropic because it consists of anisotropic particles or molecules, or because the constituent particles here assume a uniform orientation; whether the anisotropy of the Q disc signifies a peculiar *substance* or a special *condition*. Ranvier (73) interprets the anisotropy of the Q disc as the result of a

uniform arrangement of the ultramicroscopic constituent particles. He states that monorefringent bodies become birefringent by virtue of a modification in their molecular state without change of composition; that a body becomes birefringent when its molecules are oriented in one and the same axial direction. Such interpretation agrees with that of von Ebner (23) who attributes double refraction in tissues to internal stresses, producing a "photoelastic" effect. Moreover, Ambronn and Frey (1) and Freundlich (26) have shown that a flowing liquid with minute *isotropic* rods, oriented with their long axes parallel to the direction of the current, appears doubly refractive when viewed in plane polarized light transmitted perpendicularly. This condition suggests an analogy with that in the case of glass and rubber under stress. Glass is normally isotropic; when put under pressure it becomes anisotropic. Rubber is ordinarily isotropic; when stretched it becomes anisotropic. In a similar way, possibly, when the muscle fiber is at rest the molecules of the Q disc become aligned along the axis of stress, namely, parallel to the long axis of the fiber. This would follow from the relatively more fluid condition of the Q disc. When the fiber contracts or returns to rest the regular alignment of the molecules is disturbed due to the movement of the dark substance of Q from M to Z, and in the reverse direction. That this movement of the dark and stainable substance involves the salts of Q has been demonstrated by Menten (66). Discrepancies in claims regarding the anisotropic nature of the contraction bands may have their explanation in the varying degrees of disalignment and realignment of the ultramicroscopic particles of these bands at the various stages of their formation.

Muralt and Edsall (69) conclude that the muscle globulin, which contains a highly anisotropic protein (myosin), is the cause of the birefringence of the living intact muscle fiber. The anisotropy of this globulin is revealed in a current. The "double refraction of flow" in colloids, as in molecular solutions, is said to be due largely to the orientation of *anisotropic* granules. They suggest that these "anisotropic" myosin granules correspond to Bruke's "Disdiaklasten," hypothetical submicroscopic uniaxial positive doubly refractive crystals, which under the conditions of uniform orientation were held to produce the anistropy of the Q disc But the implied assumption that myosin is limited to the Q disc is unwarranted. Since myosin is presumably present throughout the entire fibril, in the J disc as well as the Q disc, the anisotropy of the Q disc cannot be due exclusively to the presence of this protein. However, assuming that the myosin granules are actually anisotropic, the

difference in optical condition and the degree of anisotropy between the J and Q discs, might possibly be the result of a difference in the extent of uniform orientation due to a difference in fluid content of these two regions. The matter of special importance in this connection concerns the fact that the stainable material that moves from Q to Z during contraction must be some substance other than myosin.

A recent and very complete study of the optical properties of striped muscle was made by Bruno (7). His summary of the literature emphasizes the confusion and the contradictory claims concerning the nature of several striations.

> According to Valentin (93), Rouget (80) and Hürthle (37) the muscular fiber as a whole is doubly refractive. Floegel (25) and Meigs (64) regard the entire contractile segment except the telophragma as doubly refractive. Brücke (6), Engelmann (24), Ranvier (73), Schipiloff and Danilewsky (84), and Heidenhain (33) claim that the double refractivity resides only in the dark disc. The dark disc (Q) is held by almost all to be anisotropic; only Schafer (83) claims that it is isotropic. The telophragma likewise is held by most to be anisotropic; only Ranvier and Vlés (96) question this interpretation. The stripe of Hensen, according to Engelmann, is not visible when the fiber is observed under polarized light; Floegel considers it doubly refractive, but in less degree than the dark disc; Rollett (76) considers it to be singly refractive; Vlés describes it as "pseudoisotropic." Regarding the nature of the double refractivity, Brücke, Hermann (35), Schipiloff and Danilewsky, Danilewsky (16), Nasse (70, 71), Soli (87), and Diamare (19) claim that this is not a natural property but that it is due to the anisotropic structure of the muscular substances. Rouget, Valentin and in part Vlés are of the opinion that the alleged anisotropic nature of the muscular substance is deceptive because of the refraction and diffraction in polarization. In the opinion of Engelmann, Ranvier, von Ebner (23), Heidenhain, and Dubois (21) the anisotropic nature of all of the contractile tissues is simply accidental. In striped muscle the phenomenon of anisotrophy is said to be dependent on the intimate constitution of a tissue whose constituent parts are in a state of constant tension.

Bruno (7) concludes that the striped muscle fibers of vertebrates and arthropods consist of discs respectively of similar tinctorial but of opposite optical properties. He agrees with Dimare that the clear disc of arthropod muscle viewed in transmitted light is monorefringent, that of vertebrate muscle birefringent.* Bruno concludes further that the optical property of the telophragma (Z membrane) depends upon the variable optical condition of the clear (J) disc; that the mesophragma (M membrane) differs optically from the telophragma; that the stripe of Hensen (H disc) like the telophragma possesses diverse optical properties in different animals; and that the stripe of Engelmann (N disc, ac-

* In view of this difference my argument regarding the relation between anisotropy and the greater fluidity of the Q disc can be considered to have strict applicability only in the case of arthropod muscle.

cessory stripe) probably has no real existence but corresponds to a shadow projected from the telophragma into the adjacent clear disc, or to a series of interfibrillar granules.

Physical characteristics of the sarcomere. The sarcomere is essentially a sac containing a mixture of colloids and crystalloids. The confining membrane consists of a telophragma terminally and a portion of the sarcostylic membrane laterally. The crystalloids are largely segregated in the Q disc. Chemically, these include potassium, chlorides and phosphates (60, 66). When fresh myofibrils of insect wing muscle are placed in distilled water, they swell and become beaded. The constrictions are produced by the relatively inextensible telophragmata. The unbeaded condition can be temporarily restored by immersion in hypertonic salt solution.

In certain theories of contraction it has been assumed that the J disc is relatively more fluid. However, when placed in alcohol of higher grades the fibrils suffer relatively more dehydration at the level of the Q discs, imposing in consequence a beaded structure. The telophragma spans the widest portion of the "bead"; it maintains a straight, unwrinkled course, again indicating a relatively rigid structure. When placed in a hypertonic salt solution the fibril shows constrictions at the level of the Q discs. When put under tension the fibrils stretch relatively more in the regions of Q. These facts indicate a more fluid condition of Q as contrasted with J. This accounts for the more ready rearrangement of the granules, thus producing anisotropy.

Sarcoplasmic constituents. The perinuclear sarcoplasm contains a Golgi internal reticular apparatus, generally closely applied to the nuclear membrane. Almost invariably, the Golgi net appears as two moieties placed at opposite poles of the nucleus (Beams, 2). This region contains also an aggregation of mitochondria, granular, bacillary and filamentous in form (Luna 59; Cowdry, 14); and other lipoid bodies, the liposomes (Bullard 7, 8). Similar mitochondria and liposomes occur sparingly throughout the interfibrillar sarcoplasm (Cowdry). Holgrem (36) describes also a widespread trophospongium. The true status of this canalicular system is still undetermined.

In the striped muscle of insects, especially the wing (thoracic) muscles, occur also so-called J and Q granules, the sarcosomes (fig. 2). These are interfibrillar spheroidal and oval granules segregated near the mid-levels of the J and Q discs. They are at least in part lipoidal in nature, and may represent large liposomes. Only typically small granular liposomes, indiscriminately scattered, occur in mammalian muscle. The fact that the J and Q granules of insect muscle do not pass across the telophragma

demonstrates that this membrane is non-fenestrated. The mesophragma offers no such effective barrier against the migration of these granules indicating either a fenestration or extreme tenuity or fragility. Bullard (8) regards the fat droplets of muscle fibers as reserve foodstuffs. In heart muscle the phospholipines occur as mitochondria.

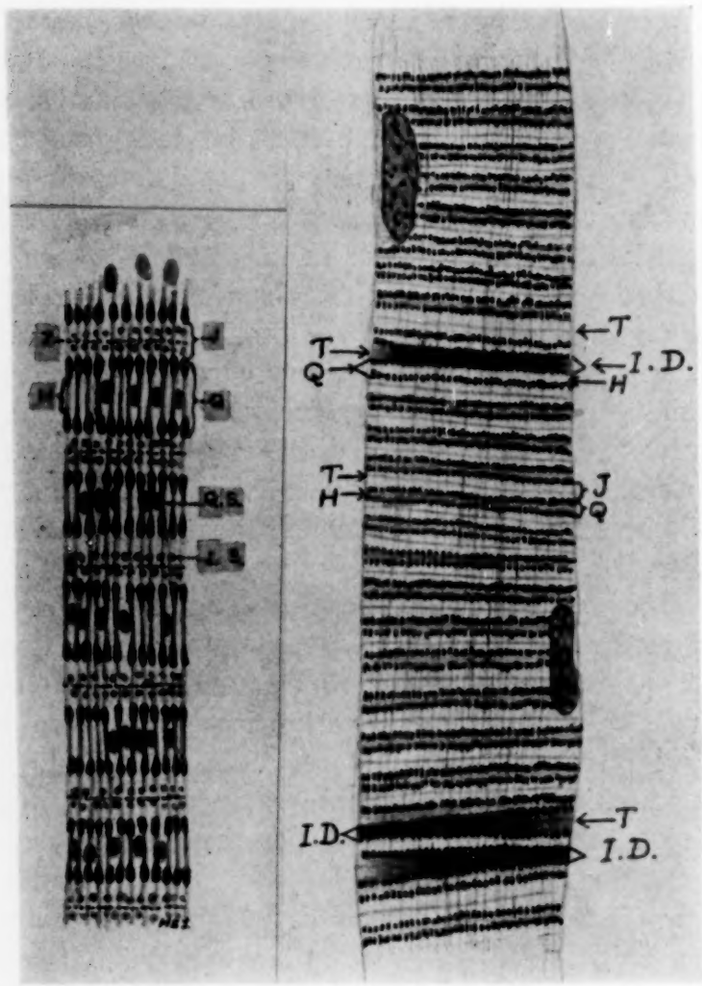

Fig. 2. Stained longitudinal section of portion of a wing muscle fiber of the "praying mantis," at mid-phase of contraction. Magnification 1600 diameters.

Fig. 3. Stained longitudinal section of fiber of human leg muscle, showing intercalated discs (*I.D.*). *T*, telophragma. Magnification 1500 diameters. (Anat. Rec., 1919, xvi, 203.)

The contracting sarcomere. The first indication of contraction is the appearance of the H disc, a light area bisecting the Q disc. As contraction proceeds the H discs widen, the opposite halves of the Q discs mov-

ing distally against the terminal telophragmata of each sarcomere. These same steps are repeated during relaxation, but in reverse order. It is not always possible to determine in sections in which direction the process is going. Moreover, fixed waves of contraction or relaxation give a false impression regarding the possible rapidity of the contraction process.

Grossly, contraction of the muscle fibril is a matter of the decrease in length and increase in width of the constituent sarcomeres. A cytologic description of the process must include at least three successive sarcomeres, for the end result is a *contraction band* consisting of a telophragma

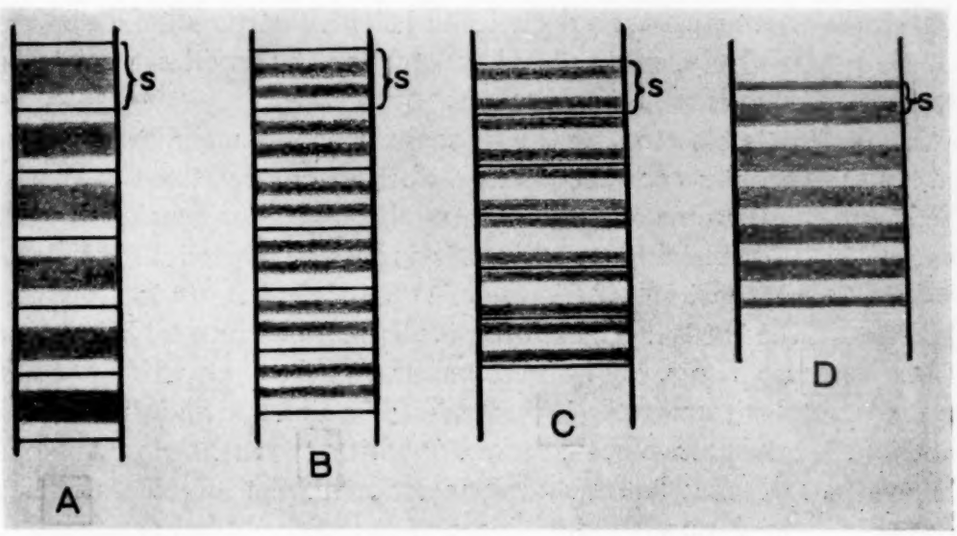

Fig. 4. Diagrams illustrating four successive stages in contraction of a striped muscle fibril. The dark stripes in *A* represent the Q discs, in *D* the dark stripes are contraction bands. *S* represents one inokomma.

flanked on either side by a half of Q of adjacent sarcomeres (fig. 4, D). Beginning with the middle one of the three sarcomeres here under consideration, the Q disc becomes bisected and the resulting halves move in opposite directions to meet at the terminal telophragmata similar halves of the Q discs of the adjacent sarcomeres (fig. 4, B and C). The result is a compact, thin, dark deeply staining disc, the contraction band, bisected by a telophragma (fig. 4, D). In the migration process of the halves of the Q discs across the J discs, the N discs become involved and presumably form part of the contraction band. The N disc is apparently interpretable as a lagging portion of the Q substance in the process of relaxation. It is not properly explained as an interfibrillar

constituent. It is not the result of the aggregation of J granules as claimed by Retzius (75). Holgrem (36), Heidenhain (31) and D'Ancona (15) conclude that the contraction band is an aggregation of sarcosomes. D'Ancona claims further that Q and J are fixation artifacts, and that the cross striation is not the cause but the result of contraction. These claims lack adequate cytologic evidence.

In the contraction process, the stainable Q substance thins out progressively in the transverse plane; the combined width of the two halves of Q during contraction becomes increasingly less than the width of the undivided Q disc of the uncontracted fibril. This result follows from the progressive increase in the diameter of the sarcomeres during contraction concomitant with their decrease in height. Meanwhile the H disc widens progressively, until at full contraction it fills all the space between successive contraction bands.

The H disc is bisected by the mesophragma (M membrane). This membrane becomes conspicuous in a fibril in contraction that has at the same time become secondarily stretched. Under usual conditions the relatively delicate mesophragma is unrecognizable; its extreme tenuity is beyond the limits of microscopic vision. Moreover, in the resting condition of the fibrils this membrane is obscured by the Q substance. It is only when the H disc has appeared, and when the contracting fiber has become stretched, thus releasing transverse tension and allowing a thickening of the mesophragma, that it enters within the range of microscopic observation. It may be concluded, accordingly, that a mesophragma is present in striped muscle though it may not ordinarily be demonstrable.

Certain theories of muscle contraction assume that J substance passes into the Q disc during the contraction process (Schaefer, 82). If such assumption were correct, the Q discs should become lighter along the J borders. On the contrary, they become lighter first in the mid-portion. This clearing mid-portion becomes the H disc. The H disc results from the escape of stainable Q substance. Nor is fluid apparently moved in either direction between J and Q. What moves is the dark substance of Q, probably the salts. The water apparently remains. Just as in the relaxed fiber, the area of M (Q disc) contains relatively more fluid than the area about Z (J disc), so in the contracting fiber the H disc (area about M) contains relatively more fluid. A superimposed tension during the contraction process affects principally the H area, producing a lengthening and a decrease in diameter; a beading, with the beads represented by the forming contraction bands at the Z levels. Moreover, the fibrils are apparently less resistant to a stretching force during contraction than

when in the relaxed condition. When the resting fiber is mechanically extended the Q substance divides medially and the H disc appears.

It becomes necessary to distinguish between the clear disc of the relaxed fiber (J disc) and that of the contraction phase (H disc); also between the H discs of a fiber in the early stages of contraction and those in the final stages. The clear discs of the later stages of contraction are actually J substance, an actual reversal of Q and J substances has occurred during contraction.

Regardless of the fact whether the myofibril is relatively slender as in vertebrates generally, or coarse as in the case of the sarcostyle of the wing muscles of diptera, hymenoptera and coleoptera, or lamellar as in the case of orthoptera (e.g., mantis), the fundamental striping is the same, consisting of Q, J and Z (45, 46, 47). This fundamental structural correspondence would seem legitimately to permit the conclusion that the physico-chemical explanation of contraction is essentially identical for all varieties of striate muscle, whether invertebrate or vertebrate, whether wing muscle or leg muscle.

Theories of muscle contraction. Several theories of muscle contraction, purporting to be based on detailed structural changes, ascribe fundamental importance to intrafibrillar canaliculi. Schaefer (82) first called attention to the evidence for such "pores." He describes them in some detail in the sarcostyle of the beetle (Dytiscus). He claims that they are open only towards the J discs, closed within the Q discs. His theory of contraction is based on the assumption that fluids pass from J through these canaliculi into Q. This is supposed to cause a swelling of Q and result in contraction without reversal of striae. Aside from the facts that a reversal of striae (strata) does actually occur during functional contraction, and that the Q discs first become clear medially instead of terminally as required by Schaefer's hypothesis, and that the myofibrils actually have a knobbed appearance in the Q region terminally during mid-phase of contraction (fig. 2), indicating movement of dark substance towards Z, vacuolated (canaliculated) sarcostyles occur only in insect wing muscle, and here only occasionally. I have seen the condition only very rarely in the wing muscle of the wasp and the elater. I feel convinced that the alleged "canaliculi" and "pores" represent artifacts.

A similar criticism is warranted with regard to the sarcostyle which Schaefer considers in contracted condition. This sarcostyle is beaded, the beads consisting of the swollen sarcomeres, the constrictions lying at the levels of the Z membranes. Such a sarcostyle is properly interpreted as altered by endosmosis. The various techniques employed

for staining of such sarcostyles are more or less hypotonic; treated subsequently with a hypertonic salt solution, such stained fibers lose their beaded condition. Normally contracting living sarcostyles are not appreciably beaded. Sarcostyles, which become beaded by treatment with hypotonic fluid, show shortened and thickened sarcomeres; such fibrils are contracted, but the contraction is an osmotic phenomenon superimposed on a dead fibril. Functional contraction is an entirely different matter; the fibril remains unbeaded and shows a reversal of striations.

The foregoing leads logically to a consideration of the question whether the sarcostyle has a membrane. Meigs (64) denies the presence of a sarcostylic membrane on the sarcostyles of the wing muscle of the house fly. However, the various reactions of the sarcostyle to hypo- and hypertonic solutions clearly demonstrate the presence on the sarcostyle of a pellicle with the properties of a semi-permeable membrane.

An outline of a theory of contraction based upon the facts as described in this paper, involving essentially the movement of Q substance (electrolytes) through J substance (colloids) to terminal telophragmata, is outside the scope of this paper. However, such an explanation, in physico-chemical terms, is attempted in my papers, Studies on Striped Muscle Nos. 6 and 7 (46, 47, 48). Here are discussed also the similar surface tension theories of Lillie (58) and the electrocapillary theory of Prenant, Bouin and Maillard (72). The more recent explanation of Clark (13) is also conceived in physico-chemical terms. She regards the Q band as composed of substances in liquid-crystal state. She accounts for the mechanical shortening of muscle contraction by the hypothesis that the substance in the Q band passes abruptly from liquid crystal to solid crystal form as a result of increase of acidity due to lactic acid formation. This explanation recognizes by implication the presence of salts in the Q band; but it disregards the phenomenon of reversal of striation in contraction and to that extent is no more adequate than the theories of Schaefer and Lillie.

Anisotropy during contraction. Merkel (65) and Rollett (78) identified an anisotropic "substance" with the material which gives the dark color to the "dim" disc of striped muscle when viewed fresh with ordinary light. These investigators noted a reversal of striation during contraction resulting in the formation of contraction bands. They inferred that the contraction band was composed of the anisotropic substance of dim disc of the relaxed fiber. Tourneux (92) and Rutherford (81) also accept this interpretation of the contraction bands. It is significant,

however, that none of these investigators gives an illustration of the contraction bands as seen in polarized light under crossed nicols. It has apparently been impossible to demonstrate an actual reversal of striae as concerns the anisotropic condition. Engelmann (24) and van Gehuchten (94, 95) actually show, as judged from their illustrations, that the anisotropic "substance" does not change its location during contraction. Schaefer (82) claims to have demonstrated the same fact by use by Rollett's gold-chloride technique. But Schaefer, as a matter of fact, confused an artificially swollen condition of the fiber, resulting from the action of the hypotonic formic-acid-water solution of this technique, with a condition of functional contraction. The assumption that the darker quality of the dim disc of striped fibers is wholly the result of the segregation of anisotropic materials within the general limits of this disc under certain conditions has led to much confusion, and has introduced an element of unnecessary difficulty in interpreting the morphologic changes of striped muscle fibers during contraction.

This undue emphasis upon the presence of anisotropic "materials" in contractile tissue has influenced some of the leading theories of muscle contraction culminating in Engelmann's (24) hypothesis that contraction is essentially a matter of the absorption of isotropic by the anisotropic material. But the presence and segregation of specific anisotropic materials cannot be held responsible for the dark color of the "dim" disc in unstained fibers; nor are such alleged anisotropic materials properly interpreted as the substance which takes the stain in the dark disc in stained preparations. The stainable dark disc is the result of a substance apart from, and in addition to, the alleged anisotropic materials. The striping of the wing-muscle sarcostyles of the wasp and other insects is quite as definite, both in the unstained and stained condition, as in the leg muscle; yet only in the latter material can the stratification of the anisotropic "substance" under certain conditions be clearly demonstrated with the micropolariscope, while in the former it is present in so small amount as to be practically impossible of demonstration in the single sarcostyle. The substance that moves during contraction in the formation of the contraction band is this additional deeply staining material. I (47) have convinced myself by a study of the leg muscle of the wasp that the anisotropic "material" (condition) as segregated in the relaxed fiber is not shifted into the light (J) disc during contraction. It is also clear, however, that the anisotropy of the fiber is greatly lessened by contraction. This latter fact is emphasized also by Meigs (64). This notoriously feebler anisotropy of the contracted fiber may result from a

disturbance during contraction of the more regularly oriented particles of the extended fiber.

The most favorable conditions for detecting the anisotropic striae are fixation in alcohol and mounting in glycerin. Previous coagulation of the sarcoplasm with alcoholic fixation gives to the anisotropic condition a more stable form and a sharper definition. But it is also quite probable that fixation of itself brings about a sharper segregation than is characteristic of the normal living sarcoplasm. Granting, however, a more or less sharp segregation of the anisotropic state, as we are in fact compelled to do for living fibers of insect leg muscle under certain normal conditions, the question remains as to why those segregations correspond in general with the limits of the dark (Q) disc. The answer to this question follows from the fundamental physical constitution of the Q disc. It was shown above that the action of alcoholic fixation and the mechanical factors of tension both indicate the relatively more fluid nature of the substance of the dark disc. Dehydration was relatively greater, producing a constriction, in the Q disc, and tension resulted first in an elongation of the Q disc. Since anisotropy is the result of the orientation of the molecules of a substance in the same axial direction (Ranvier), we should expect this optical condition accentuated in that portion of the fiber, namely, the more fluid Q portion, where the molecules or particles are relatively more free to orient themselves (during relaxation) in the axis of stress, that is, parallel to the longitudinal axis. The corollaries to this explanation are: 1, that the disturbance of this identical orientation of the particles of the Q disc during contraction results in a greatly lessened or a vitiated anisotropy of the fiber; and 2, those relaxed fibers which are in an apparently inactive condition as regards especially stratification may be such in which the particles of the Q disc have not yet become reoriented in the longitudinal axis immediately after contraction.

The reversal of striae, however, as seen in unstained and especially in stained fibers is a function of a substance unrelated fundamentally to the anisotropic phenomenon, and is the result of the change in position during contraction of some other than a specifically anisotropic substance from the vicinity of the mesophragmata to the vicinity of the telophragmata. Reversal of striation, as regards a deeply staining substance, and contraction are coincident phenomena. The phenomenon of reversal of striae, and the chemical constitution of the deeply staining substance involved, are conditions obviously more closely related to contraction than the very widespread (organically and inorganically) condition of anisotropy. Any final and adequate explanation

of contraction must be able to embrace the fundamental morphologic datum of an intrasarcostylic reversal of striation as regards the deeply staining substance of the sarcoplasm during contraction.

The intercalated discs. The structure of the contraction band of striated muscle gives the basis for a consideration of the intercalated discs of cardiac muscle. Heart muscle is substantially like voluntary striped muscle. Its Q discs are relatively less conspicuous, the telophragmata are relatively more robust and N discs are apparently lacking; but it consists essentially of myofibrils composed of a series of sarcomeres in every respect like skeletal muscle. The syncytial character of definitive cardiac muscle, in contrast with the cellular condition of skeletal muscle need not detain us.

The outstanding characteristic of cardiac muscle is the occurrence of intercalated discs. These have been variously interpreted. Heidenhain (31) regards them as zones of growth in post-fetal life; Schweigger-Seidel (85) thought they represented intercellular cement lines; Marceau (62) interpreted them as tendons. Benda (5) and Renaut (74) held ideas very similar to that of Marceau; they attached only mechanical significance to these discs. Dietrichs (18) considers them coördinating and supporting mechanisms for bundles of fibrils, and Witte (97) interprets them as "strengthening bands." According to von Ebner (23) they signify no more than the products of an irregular, abnormal or agonal local contraction. Kato (52) advances the novel claim that they serve as a mechanism for the transport of nutritive materials.

The cytologic evidence contradicts all such interpretations. These discs are explainable as "modified irreversible contraction bands" (48, 49, 50, 51). If such interpretation is correct, it should be possible under certain conditions to find comparable structures in skeletal muscle. Such expectation has been realized in a specimen of human leg muscle (fig. 3) from a young athlete (44).

Intercalated discs of heart muscle occur in the form of straight bands, stepped bands, and serrated structures. The latter predominate in, and are characteristic of hypertrophy. Atrophy is characterized by a compact "comb" form. The "teeth" of the "comb" represent modified portions of the involved myofibrils. During early developmental stages only band forms occur, in sections passing in part completely, in part for variable distances, across the surface of the fibers. Subsequently, probably as the result of oblique tensions due to the profuse branching of the cardiac trabeculae, the bands become modified into step-like forms. They may have a supernuclear position. These so-

called discs are in reality bands; they are placed superficially; they penetrate to variable depths but never completely span the fiber. They occur as superficial crescents, rings or spirals. The elementary unit of the intercalated disc is a portion of the myofibril in the immediate vicinity of the telophragma. This interpretation is confirmed by the more recent work of Galiano (27) by use of the Hortega technique for neuroglia. By use of this same technique D'Ancona (15) demonstrates very clearly that these discs frequently separate contracted from non-contracted areas. In their entirety these discs include also the the interfibrillar sarcoplasm. The serration of the discs of hypertrophied muscle results from the longitudinal splitting of the myofibrils involved in hypertrophy. The sides of the serrations represent interconnecting portions of the telophragma. The irregularities of the serrations may be interpreted as the result of unequal stresses in adjacent new-formed fibrils.

The original and prevailing simple band-form of normal intercalated disc is either bisected by a telophragma, or bounded on one side by this membrane or bounded on both sides by a telophragma. In the form of a "band" bisected by the telophragma it corresponds exactly to a contraction band as above described. Assuming a condition of inability to reverse the contraction process, the contraction band becomes at once a simple intercalated disc. The factor involved is presumably hyperacidity, perhaps excess of lactic acid, effecting an irreversible coagulation.

It will be recalled that the contraction band is essentially a double structure, consisting of opposite halves of successive Q discs fused about a telophragma. Under the assumed conditions, such as effect irreversible coagulation (contraction), it is quite conceivable that half of the contraction band might be capable of reversible action while the other half remained coagulated. This condition would result in an intercalated disc bounded on one side by a telophragma. Assuming further that adjacent halves of successive contraction bands remained irreversible, it is conceivable that following relative contractile inactivity of such local areas, the two halves would fuse into a single structure, producing thus the occasional type of disc bounded on both sides by a telophragma. The interpretation of intercalated discs as modified irreversible contraction bands seems fully to meet all requirements for explanation of the various morphologic types.

The status of the myofibril. Thus far it has been assumed that myofibrils are actual morphologic entities. Since such entity has been

denied or questioned by certain workers, it becomes necessary to discuss this matter and to prove the actual existence of muscle fibrils. If myofibrils are fixation or rigor mortis artifacts, then the foregoing description of morphologic changes in non-existent structures during contraction becomes fatuous. Lewis and Lewis (57), M. R. Lewis (53, 54, 55), W. H. Lewis (56), and Goss (29, 30) record their inability to observe myofibrillae in living striped muscle cells in tissue cultures, and incline to an interpretation of fibrils as fixation artifacts. In proof of the actual existence of fibrils and cross striations (discs) I would offer first a personal observation under high magnification of striped muscle in the tail of young frog and salamander tadpoles. One who takes the trouble to have this experience will have no further doubt that fibrils and cross stripes are real structures in living muscle. The further step of actually watching the "transposition of stripes" during contraction has not yet been taken personally to my complete satisfaction. However, seeing how perfectly the fixed structure corresponds with that of the living fibers, one need entertain no further doubt of the actual reversal of striation in contraction as revealed by properly fixed and stained sections. The fact that myofibrils are not discernible in living muscle cells under certain conditions is not proof that they are non-existent; they may be too close in refractive index to the adjacent myoplasm to become visible under the microscope. Mechanical tensions or coagulation may be necessary to produce the contrast required for clear visibility.

Carr's (11) demonstration of the continuity of myofibrils with tendon fibrils in skeletal muscle of the fetal pig gives additional support to the interpretation of myofibrils as genuine morphologic elements. Moreover, the mode of fibril formation in the skeletal muscle of teleosts supplies unequivocal evidence of the structural entity of myofibrils. In the rain-bow trout, for example (32, 48), there appears close to the nucleus in the early myoblasts, a single, relatively coarse fibril. This primary fibril is oriented parallel to the long axis of the cell. By longitudinal fission it divides into four secondary fibrils. Subsequent radial and concentric longitudinal divisions contribute the groups (Kölliker's columns) of definitive myofibrils.

SUMMARY

In essence contraction is a matter of the equal division of the dark, stainable materials of the Q discs and the movement of the resultant semi-discs in opposite direction against the limiting telophragmata of a sarcomere. Relaxation is the reversal of this process, the halves of re-

spective Q discs returning to their originally bisecting meosphragmata. Thus, two limiting membranes (terminal Z and median M), a laterally confining pellicle (sarcostylic membrane) and a dark substance (salts of Q) moving through a colloidal clear substance, are the only essential morphologic elements involved in contraction. How this shuttle-like movement of a dark substance between successive telophragmata effects contraction and relaxation would seem to be the central problem of the physiologist in his efforts to formulate an adequate explanation of muscle contraction.

BIBLIOGRAPHY

(1) Ambronn, H. and A. Frey. 1926. Das Polarisationsmikroskop. Leipzig. (Cited from Takahashi and Rawlins, Science, Jan. 6, 1933.)

(2) Beams, H. W. 1929. Studies on the "Golgi apparatus" of insect muscle. Anat. Rec., xlii, 323.

(3) Bell, E. T. 1911. The interstitial granules of striated muscle and their relation to nutrition. Internat. Monatschr. f. Anat. u. Physiol., xxviii.

(4) Benda, C. 1899. Weitere Mittheilugen über der Mitochondria. Arch. f. Anat. Physiol., Abt., 376.

(5) Benda, C. 1902. Ueber den feineren Bau der glatten Muskelfaser des Menschen. Verhandl. d. Anat. Ges. zu Halle.

(6) Brücke, E. 1858. Untersuchungen über den Bau der Muskelfasern mit Hilfe der polarisierten Lichtes. Wiener Denkschr. mat.-nat. Klasse, xv, 69.

(7) Bruno, G. 1930. Proprietà ottiche delle fibre muscolari striate dei vertebrati e degli artropodi. Arch. Ital. di Anat. e. Embr., xxvii, 613 (136 references).

(8) Bullard, H. H. 1912. On the interstitial granules and fat droplets of striated muscle. Amer. Journ. Anat., xiv, 1.

(9) Bullard, H. H. 1916. On the occurrence and physiological significance of fat in the normal myocardium and atrio-ventricular system (bundle of His), interstitial granules (mitochondria) and phospholipines in cardiac muscle. Amer. Journ. Anat., xix, 1

(10) Cameron, J. 1917. The histogenesis of vertebrate striated muscle, including a contribution to our knowledge regarding the structure and functions of the cell-nucleus. Trans. Roy. Soc. of Canada, xi, 81.

(11) Carr, R. W. 1931. Muscle-tendon attachment in the striated muscles of the fetal pig; demonstration of the sarcolemma by electric stimulation. Amer. Journ. Anat., xlix, 1.

(12) Chittenden, R. H. 1883. Histochemische untersuchunger über das Sarkolemm und einige verwandte membranen. Untersuch. Physiol. Inst. U. Heidelberg., iii, 171.

(13) Clark, J. H. 1927. A theory of muscle contraction with x-ray diffraction patterns from relaxed and contracted muscles. Amer. Journ. Physiol., lxxxii, 181.

(14) Cowdry, E. V. 1918. The mitochondrial constituents of protoplasm. Pub. Carnegie Inst. Wash., viii, 39.

(15) D'Ancona, U. 1929. Contributo a una revisione delle nostre conoscenze sulla morfologia della fibra muscolare striata. Protoplasma, x, 177 (196 references).
(16) Danilewsky, A. 1882. Ueber die Abhängigkeit der Kontractionart der Muskeln von den Mengenverthältnissen einiger ihre Bestandtheile. Zeitschr. f. physiol. Chem., vii.
(17) Davenport, H. K. and S. W. Ranson. 1930. Contracture resulting from tenotomy. Arch. Surg., xxi, 1.
(18) Dietrich, A. 1910. Die Elemente des Herzmuskels. Fischer, Jena, 1.
(19) Dimare, V. 1923. Sui corpi mielinici del sangue e die muscoli striati e sull anisotropismo della fibra striata. Riv. di. Biol., vi, 6.
(20) Dimare, V. 1927. Indogini ottiche e interpretazioni chimicofisiche di strutture nella fibra musculare striata. Rend. R. Acad. Sc. Napoli., Ser 3ª, xxxv (Cited from D'Ancona).
(21) Dubois, R. 1909. Recherches sur la pourpre et les pigments animaux. Arch. Zool. Exper. (Cited from Bruno).
(22) Duesberg, J. 1909. Les chondriosomes des cellules embryonaires du poulet, et leur rôle dans le genèse des myofibrilles, avec quelques observations sur le dévéloppement des fibres musculaires striées. Arch. f. Zellforschung, iv, 602 (154 references).
(23) Ebner, V. 1882 Untersuchungen über die Ursachen der anisotropie organischer Substanzen. Leipzig.
(24) Engelmann, T. W. 1878. Neue Untersuchungen über die mikroscopischen vorgange bei der Muskelkontraktion. Pflüger's Arch., xviii.
(25) Floegel, J. H. L. 1872. Ueber die quergestrieften Muskeln der Milben. Arch. f. mikr. Anat., viii.
(26) Freundlich, H. Colloid and capillary chemistry. E. P. Dutton & Co., New York. (Cited from Takahashi and Rawlins, Science, Jan. 6, 1933.)
(27) Galiano, E. F. 1926. Sobre la estructura y la significacion funcional de las piegas intercalares del corazion. Bol. Real. Sociedad Española de Hist. Nat., xxvi, 227.
(28) Goddard, M. 1839. On the polarization of light by living animals. Phil. Mag., ser. iii, xv.
(29) Goss, C. M. 1931. "Slow-motion" cinematographs of the contraction of single cardiac muscle cells. Proc. Soc. Exp. Biol. and Med., xxix, 292.
(30) Goss, C. M. 1932. The formation of cross-striations in cultures of embryonic heart muscle. Arch. f. exp. Zellforschung., xii, 233.
(31) Heidenhain, M. 1911. Plasma and Zelle. G. Fischer, Jena.
(32) Heidenhain, M. 1913. Ueber die Entstehung der quergestreiften Muskelsubstanz. bei der Forelle. Beiträge zur Teilkörpertheore. II. Arch. f. mikr. Anat., lxxxiii, 427.
(33) Heidenhain, R. 1861. Zur Frage nach der Form der kontraktilen Fasergellen. Studien d. phys. Inst. zu Breslau. H 1 (Cited from Bruno).
(34) Hensen, V. 1868. Ueber ein neues Strukturverhaltnis der quergestreiften Muskelfaser. Arb. d. Kieler Physiol. Inst.
(35) Hermann, M. 1880. Ueber das verhalten der optischen Constanten des Muskels. Pflüger's Arch., xxii.

(36) HOLGREM, E. 1907. Ueber die Sarkoplasmakörner quergestreifter Muskelfasern. Anat. Anz., xxxi, 609.

(37) HÜRTHLE, K. 1907. Ueber die Struktur der quergestreiften Muskels in ruhenden und tätigen Zustande und über seinen aggregätzustand. Biol. Centralbl., xxvii.

(38) JORDAN, H. E. 1911. The structure of heart muscle of the humming bird, with special reference to the intercalated discs. Anat. Rec., v, 517.

(39) JORDAN, H. E. 1912. The intercalated discs of hypertrophied heart muscle. Anat. Rec., vi, 357.

(40) JORDAN, H. E. 1916. A comparative microscopic study of cardiac and skeletal muscle of Limulus. Anat. Rec., x, 1.

(41) JORDAN, H. E. 1916. The microscopic structure of the leg muscle of the sea-spider, Anoplodactylus lentus. Anat. Rec., x, 493.

(42) JORDAN, H. E. 1917. The microscopic structure of striped muscle of Limulus. Pub. 251, Carnegie Inst. Wash., 273.

(43) JORDAN, H. E. 1917. Studies on striped muscle structure. III. The comparative histology of cardiac and skeletal muscle of scorpion. Anat. Rec., xiii, 1.

(44) JORDAN, H. E. 1919. Studies on striped muscle structure. IV. Intercalated discs in voluntary striped muscle. Anat. Rec., xvi, 203.

(45) JORDAN, H. E. 1919. Studies on striped muscle structure. V. The comparative histology of the leg and wing muscles of the mantis, with special reference to the N-discs and the Sarcosomes. Anat. Rec., xvi, 217.

(46) JORDAN, H. E. 1920. Studies on striped muscle structure. VI. The comparative histology of the leg and wing muscle of the wasp, with special reference to the phenomenon of stripe reversal during contraction and to the genetic relation between contraction bands and intercalated discs. Amer. Journ. Anat., xxvii, 1 (32 references).

(47) JORDAN, H. E. 1920. Studies on striped muscle structure. VII. The development of the sarcostyle of the wing muscle of the wasp, with a consideration of the physico-chemical basis of contraction. Anat. Rec., xix, 97.

(48) JORDAN, H. E. 1930. A text book of histology. D. Appleton & Co., New York, 100.

(49) JORDAN, H. E. AND J. B. BANKS. 1917. A study of the intercalated discs of the heart of the beef. Amer. Journ. Anat., xxii, 285.

(50) JORDAN, H. E. AND J. BARDIN. 1913. The relation of the intercalated discs to the so-called "segmentation" and "fragmentation" of heart muscle. Anat. Anz., xliii, 612.

(51) JORDAN, H. E. AND K. B. STEELE. 1912. A comparative microscopic study of the intercalated discs of vertebrate heart muscle. Amer. Journ. Anat., xiii, 151.

(52) KATO, Y. 1928. Experimentelle Untersuchungen an den Herzmuskelfasern. I. Über die Schaltstücke der Herzmuskelfasern, insbesondere ihre Bedentung. II. Über die Entstehung der Schaltstücke an Herzmuskelfasern beim jungen Kaninschen. Folia Anat. Jap., vi, 637.

(53) LEWIS, M. R. 1915. Rythmical contraction of the skeletal muscle tissue observed in tissue cultures. Journ. Physiol., xxxviii, 153.

(54) LEWIS, M. R. 1919. The development of cross-striations in the heart muscle of the chick embryo. Johns Hopkins Hosp. Bull., xxx, 176.
(55) LEWIS, M. R. 1919. Muscular contraction in tissue-cultures. Contributions to Embryology, no. 35; Pub. 272 Carnegie Inst. Wash.
(56) LEWIS, W. H. 1923. Cultivation of heart muscle from chick embryo (4 to 11 days old) in Locke—bouillon—dextrose medium. Anat. Rec., xxv, 111 (Proc. Amer. Assoc. Anat.).
(57) LEWIS, W. H. AND M. R. LEWIS. 1917. Behavior of cross-striated muscle in tissue cultures. Amer. Journ. Anat., xxii, 169.
(58) LILLIE, R. S. 1912. The physiological significance of the segmented structure of the striated muscle fiber. Science, xxxvi, 247.
(59) LUNA, E. 1913. Sulla importanza dei condriosomi nella genesi delle myofibrille. Arch. f. Zellforsch., ix, 458.
(60) MACALLUM, A. B. 1905. On the distribution of potassium in animal and vegetable cells. Journ. Physiol., xxxii, 95.
(61) MCDOUGALL, W. 1897. On the structure of cross-striated muscle and a suggestion as to the nature of its contraction. Journ. Anat. and Physiol., xxi, 410.
(62) MARCEAU, F. 1904. Recherches sur la structure et le developpement compare des fibres cardiaque. Ann. des. Sc. Nat. Zool., xix.
(63) MARCUS, H. 1921. Über den feineren Bau quergestreifter Muskeln. Arch. f. Zellforsch., xv, 393.
(64) MEIGS, E. B. 1908. The structure of the element of cross-striated muscle, and the changes of form which it undergoes during contraction. Zeitschr. f. Allg. Physiol., viii, 81.
(65) MEIGS, E. B. 1912. Contributions to the general physiology of smooth and striated muscle. Journ. Exp. Zool., xiii, 497.
(66) MENTEN, M. L. 1908. The distribution of fat, chlorides, phosphates, potassium and iron in striated muscle. Trans. Canadian Inst., viii, 403.
(67) MERKEL, F. 1872. Der quergestreifte Muskel. I. Das primitive Muskelelement der Arthropoden. Arch. f. mikr. Anat., viii, 244.
(68) MERKEL, F. 1873. Der quergestreifte Muskel. II. Der Contractionsvorgang im polarisierten Lichte. Schultze's Arch., ix, 293.
(69) MURALT, A. L. VON AND J. T. EDSALL. 1930. Studies in the physical chemistry of muscle globulin. III. The anisotropy of myosin and the angle of isocline. Journ. Physiol. Chem., lxxxix, 315. IV. The anisotropy of myosin and double refraction of flow. Ibid., 351.
(70) NASSE, O. 1878. Zur mikroscopischen Untersuchung des quergestreiften Muskels. Pflüger's Arch., xvii.
(71) NASSE, O. 1882. Zur Anatomie und Physiologie der quergestreiften Muskelsubstanz. Leipzig.
(72) PRENANT, A., P. BOUIN, ET L. MAILLARD. 1904. Traité d'Histologie, i, 440.
(73) RANVIER, L. 1880. Lecons d'Anatomie Général sur le système Musculaire. Paris.
(74) RENAUT, J. 1911. Note sur les disques accessoires des disques minces dans les muscles stries. Comptes rendus, lxxxv.
(75) RETZIUS, G. 1890. Muskelfibrillae und Sarkoplasma. Biol. Unters. von Retzius.
(76) ROLLETT, A. 1885. Untersuchungen über den Bau der quergestreiften Muskelfasern. I und II. Wien.

(77) Rollett, A. 1891. Über die N-Streifen (Nebenscheiben) das Sarkoplasma und die Kontraction der quergestreiften Muskelfasern. Arch. f. mikr. Anat., xxxvii, 654.

(78) Rollett, A. 1891. Untersuchungen über Kontraktion und Doppelbrechung der quergestreiften Muskelfasern. Wien.

(79) Rollett, A. 1892. Ueber die Kontraktionswellen und ihre Beziehung zu der Entzuckung bei den quergestreiften Muskelfasern. Pflüger's Arch., liv.

(80) Rouget, C. 1862. Sur les phénomènes de la polarization qui s'observant dans quelques tissus des végétaux et des animaux. Journ. de la Physiol., v.

(81) Rutherford, W. 1897. On the structure and contraction of striped muscular fiber. Journ. Anat. and Physiol., xxxi, 309.

(82) Schaefer, E. A. 1910. Essentials of histology. Longmans, Green & Co.

(83) Schaefer, J. 1873. Beiträge zur Histologie und Histogenese der quergestreiften Muskelfasern des Menschen und einiger Wirbeltiere. Wiener Sitzungsber., cii.

(84) Schipiloff, C. and A. Danilewsky. 1881. Ueber die Natur der anisotropen Substanzen des quergestreiften Muskels und ihre räumliche verteilung im Muskelbündel. Hoppe-Seyler's Zeitschr., v.

(85) Schweigger-Seidel, F. 1871. Das Herz, in S. Stricker's Handbuch der Lehre von den Geweben des Menschen und der Tiere. Leipzig.

(86) Scott, G. H. 1932. Distribution of mineral ash in striated muscle cells. Proc. Soc. Exp. Biol. and Med., xxix, 349.

(87) Soli, V. 1906. Sulla struttra delle fibre muscolari lisce dello stomaco degli uccelli. Bibl. Anat., xvii.

(88) Thulin, I. 1915. Ist die Grundmembran eine konstant vorkommende Bildung in den quergestreiften Muskelfasern? Arch. f. mikr. Anat., lxxxvi, 318.

(89) Tiegs, O. W. 1922. On the arrangement of the striations of voluntary muscle fibres in double spirals. Trans. Roy. Soc. of South Australia, xlvi, 222.

(90) Tiegs, O. W. 1923. The structure and action of "striated" muscle fibre. Trans. Roy. Soc. of South Australia, lxvii, 142.

(91) Tiegs, O. W. 1924. On the mechanism of muscular action. Australian Journ. Exper. Biol. and Med. Sci., i, 11.

(92) Tourneux, F. 1892. Sur les modifications structurales que présentent les muscles jaunes du Dytique pendant la contraction. Journ. Anat. et. Physiol. norm. et. path., Année, 28.

(93) Valentin, G. 1861. Die Untersuchungen der Pflanzen und Thiergewebe in polarisirten Lichte. Leipzig.

(94) Van Gehuchten, A. 1886. Étude sur la structure intime de la cellule musculaire striée. La Cellule, ii, 293.

(95) Van Gehuchten, A. 1888. Étude sur la structure intime de la cellule musculaire striée chez les Vertébrés. La Cellule, iv, 247.

(96) Vlés, F. 1909. Sur le valeur des stries musculaires en lumiere polarisée. C. R. Soc. Sci.

(97) Witte, L. 1919. Histogenesis of the heart muscle of the pig in relation to the appearance and development of the intercalated discs. Amer. Journ. Anat., xxv, 333.

THE MODE AND SITE OF ACTION OF STRYCHNINE IN THE NERVOUS SYSTEM

J. G. DUSSER de BARENNE

Laboratory of Neurophysiology, Yale University, New Haven, Conn.

In this review we will deal only with the action of strychnine on the nervous system so far as the cerebrospinal mechanisms are concerned, excluding from discussion its action on the central and peripheral neural mechanisms of autonomic effectors (heart, smooth muscle, glands, pigment cells).

Strychnine is the outstanding representative of the so-called *central* convulsant poisons. The typical and early symptoms of poisoning with a sufficient dose of strychnine are a marked hyper-reflexia, soon followed by generalized convulsions due to unbalanced, maximal contractions of practically all striped muscles. The attitude of the various animals during a strychnine spasm depends upon the relative strength of the various muscle groups.

In the frog, for instance, the fore limbs are in strong flexion and adduction, because in these extremities the flexor muscles are more powerful than the extensors whereas the legs are in extension because of the mechanical predominance of the extensors over the flexor muscles in these limbs. In quadruped mammals the picture is one of strong extension of all four limbs with opisthotonus, through dominance of the extensor muscles of the limbs and the vertebral column.

These strychnine spasms are of *central* origin, as demonstrated already by Magendie (1819). Upon ligation of a hindleg of a frog, with exception of the sciatic nerve, injection of strychnine into this leg is not followed by convulsions, neither of the muscles of the leg injected, nor of the other parts of the body. The same conclusion follows from other classical data, which show that the spasms are promptly abolished 1, in a limb after section of its peripheral nerves; 2, by section of its ventral roots, or 3, by destruction of the spinal cord (pithing of the animal).

The fact that typical strychnine spasms of the limb and trunk muscles still appear in the decapitate preparation proves that these convulsions are of *spinal* origin, and that the activity of prespinal centres is not significant in their occurrence. The same conclusion can be reached

from the evidence that in a frog, either with intact central nervous system or decapitated, in which the circulation has been stopped by extirpation of the heart and the lymph hearts, local application of a dilute strychnine solution about the exposed spinal cord promptly results in typical convulsions.

It should be added that for the occurrence of the spasms of the mastication muscles and other muscles of the head, the strychninization of the corresponding centres in the brainstem, is necessary.

Mode of action of strychnine. Scientific investigations of strychnine action began with Magendie, and it was early recognized that strychnine produces not only an enormous augmentation of reflex excitability, but also a marked spread of reflex activity on any stimulus entering the central nervous system. However, closer insight into its action is of relatively recent date.

L. and M. Lapicque found in 1907 that under strychnine, either applied locally to a peripheral nerve of a frog or injected into the bloodstream, the excitability of these nerves is markedly increased, expressing itself in a decrease of their chronaxie, Lapicque's time factor in excitation. Bremer and Rylant (5) have corroborated this observation for efferent and afferent nerve fibres in mammals. These facts are certainly important, but on the other hand, it must be kept in mind that the typical effects of strychnine poisoning are brought about not by the action of this drug on the peripheral nerve fibres, but by its action on the central mechanisms. We have, therefore, to look towards these centres for an explanation of its action.

In 1905 Sherrington (18b) made the important observation that small or moderate, *non*-convulsant doses of strychnine (0.08–0.1 mgm. per kilogram bodyweight in mammals) destroy the reciprocal innervation of antagonistic muscles, one of the basic principles of normal reflex activity, reverse the normal inhibition of an antagonist into contraction, simultaneous with the contraction of the agonist. This strychnine reversal, however, is certainly not a universal phenomenon. Magnus and Wolf (14) found that there are some reflexes, namely, the tonic neck-reflexes (Magnus and deKleyn), the compensatory eye-reflexes and the so-called caloric labyrinthine eye-reflexes, which are not reversed under strychnine, even under large, convulsant doses of the drug, although the same muscles that are active in these reflexes demonstrate the existence of typical central reversal when they are activated in other reflexes. Bremer (4a) and Miller (16) found that the inhibition of the extensor muscles in decerebrate rigidity, elicited by stimulation of the

anterior parts of the cerebellar cortex, is not reversed by strychnine; neither does it reverse the inhibition of the opening muscles of the mouth in the mandibular and linguo-mandibular reflex (4b). Liddell and Sherrington (13) also showed that the reflex inhibition of a myotatic stretch reflex in an extensor muscle upon pull on its flexor antagonist remains unchanged, even under convulsant doses of strychnine.

The explanation of the appearance of Sherrington's strychnine reversal in all probability is that the inhibitory stimulus applied to the afferent peripheral nerve, as is the procedure in his experiments, includes also an excitatory factor, which normally does not predominate, but under strychnine becomes dominant and blocks the inhibitory effect, whereas the absence of this reversal in the above mentioned cases is probably due to the pure inhibitory nature of the excitations in these instances, free from any excitatory component. These facts show that we cannot look upon strychnine reversal as the fundamental phenomenon which, in destroying reciprocal innervation, works havoc in the central integrating mechanisms, but that in all probability it only appears under special experimental conditions as another example of the augmented irradiation of impulses in the central nervous system under the influence of strychnine. Cushny (6) has emphasized that in the frog under convulsant doses of strychnine it is still possible, using "weak and slowly acting stimuli," to obtain normal reflex movements.

More light has been thrown upon the mode of action of strychnine on these mechanisms by recent work, especially of Bremer and Rylant (5). These investigators found that the chronaxies of the various groups of peripheral nerves and spinal reflex arcs (which normally are different, extensor-neurones having in general a larger chronaxie than flexor mechanisms) decrease markedly under strychnine and become equalized. Furthermore they observed that the group of reflexes designated by Liddell and Sherrington as "d'emblée" reflexes, in which all motor nerve fibres or better all motor units, in Sherrington's terminology, engaged in the reflex respond to the first stimuli, are markedly augmented under strychnine, with distinct shortening of the latent period. The second type of reflexes, which the work of Liddell and Sherrington has made known, their "recruitment"-reflexes which are characterized by their long latency, slow onset, long after-discharge and slow decline, show a more profound change under strychnine. These reflexes then assume the characteristics of the d'emblée reflex type, with latency greatly reduced, onset abrupt and the maximum of the contraction reached almost immediately, little or no after-discharge and disap-

pearance of the so-called secondary waves in the electromyogram between the primary electrical waves of the reflex responses. The same characteristics pertain to the electromyograms of fully developed strychnine tetani, as has been shown by many earlier electrophysiological observations on strychnine tetani (Fl. Buchanan, Boruttau, Fahrenkamp, Hoffmann, Veszi and others). These changes in reflex activity offer also a satisfactory explanation of Cushny's "start" reflex; I doubt whether this reflex actually can be regarded as a separate phenomenon, as Cushny did.

From all this experimental evidence Bremer concludes that the essential action of strychnine is a pronounced augmentation of the excitability and conductability of the various nervous structures involved, resulting in immediate, generalized and synchronized maximal reflex responses to any extraneous stimulus.

We still lack insight into the intimate nature of these functional changes. It is often stated that strychnine takes away all resistances in the various reflex arcs; this certainly is a useful allegory, but the question remains, how strychnine acts in this way.

Site of action of strychnine. We will discuss now the problem of the site of action of strychnine: Which are the nervous structures in the central nervous system, upon which strychnine acts and through which it produces the typical tetani?

We have seen already that these tetani are of central origin; the strychnine action on peripheral nerve fibres, as found by L. and M. Lapicque, Bremer and Rylant (see above) is not essential in this respect.

Stilling was one of the first to make (19) a definite statement; he localized the action of strychnine in the anterior gray matter of the cord, in the ventral horns. Hermann Meyer (15) in 1846 clearly demonstrated that strychnine tetani are reflex in origin and stated that "strychnine excites the activity of the gray matter or better the activity of the ganglionballs ('Ganglienkugeln') which make up this matter." This in itself is quite a remarkable statement in 1846, considering the state of knowledge with regard to the histology of the nervous system in those years. In my earlier strychnine papers (7b, 7c) Meyer's statements were interpreted as indicating that he looked upon the posterior gray matter as the site of action of strychnine. This is also the most common interpretation of this paper of Meyer's in the literature. On restudying his paper after so many years I am doubtful whether this interpretation is correct; it may be that Meyer had no such specification of the posterior gray matter in mind.

Houghton and Muirhead (12) advocated the view that strychnine tetani arise from the action of the drug on sensory mechanisms in the dorsal horns of the spinal cord. They applied a dilute strychnine solution with a small pipette around the arm region of the spinal cord of the frog, in which the circulation had been stopped by excision of the heart and lymph hearts, and observed not only typical tetani in the arms, but also in the legs on stimulation of the skin of the arms. On application of the poison to the distal region of the spinal cord, typical spasms were observed, in the legs as well as in the forelimbs, on cutaneous stimulation of the hindlegs.

In 1900 Verworn (20) and especially Baglioni (1), then working in Verworn's laboratory, also propagated the theory that strychninization of the sensory dorsal mechanisms of the spinal cord is responsible for the tetani. They even went so far as to proclaim strychnine as an elective poison for sensory mechanisms in the central nervous system. Baglioni came to this conclusion because he noticed the appearance of typical strychnine tetani on application of strychnine to the dorsal surface of the cord, which poisoning he assumed to be strictly localized and to remain localized. Baglioni describes his technique as follows: (1, p. 201) "The point of a needle was surrounded by a very fine layer of cotton wool and with this needle made blunt in this way and dipped into the solution of carbolic acid [or of strychnine (Ref.)] the desired part of the spinal cord was touched. As soon as a small drop of the solution touched the cord, it was taken away with cotton wool, so as to prevent spread. This was repeated so often, until the symptoms of the poisoning appeared."

The above mentioned conclusions, however, must be looked upon as incorrect. This follows from the following experimental data: Strychnine has no appreciable effect on the most definite sensory nerve cells of the body, those of the spinal ganglia (12, 7a); strictly local strychninization of the dorsal mechanisms of the spinal cord does not result in strychnine tetani, but in a complex of symptoms, the syndrome of the strychninization of the dorsal mechanisms of the spinal cord (7b, 7e, 7g). Typical strychnine tetani appear only on simultaneous strychninization of the dorsal, sensory spinal mechanisms and of the ventral, motor mechanisms (7b). Finally strychnine acts also on autonomic central mechanisms, as was found in recent injection experiments (8, p. 264).

Let us consider first of all the effect of local strychninization of the dorsal, sensory spinal mechanisms. The typical syndrome which then appears, consists of the following symptoms:

1. Paroxysmal, paresthetic disturbances, occurring without any appreciable stimulus, even after degeneration of the posterior root fibres. 2. Hyperesthesia and hyperalgesia of the skin of the so-called strychnine segment zone(s), corresponding to the spinal segment(s), the dorsal surface of which has been poisoned. 3. Hypersensitivity of the deeper structures (muscles, tendons and probably the periost), in accordance with the segment(s) strychninized. 4. a. Hyper-reflexia; b, twitchlike, irregular, abrupt contractions,[1] occurring without any appreciable internal stimulus; these contractions are still for the greatest part reflex contractions, as they are much less conspicuous after section of the posterior roots.

No true strychnine tetani are observed in this syndrome. They appear immediately when the dorsal poisoning is followed by strychninization of the ventral mechanisms, or when the combined strychnine poisoning of these two sets of mechanisms is done in the reverse order. This result is obtained in the frog, as well as in the mammal. The most convincing way of performing this experiment is the following: Exposure of the dorsal surface of the spinal cord in the arm region and in the leg region. First strictly local application of strychnine on the *dorsal* surface of the arm segments, which immediately is followed by the appearance in the frontlegs of the syndrome described above, in which no true tetani can be observed. Subsequently strychnine is applied locally on the *ventral* side of the spinal cord in the region of the segments for the hindlimbs. Then stimulation of the skin of the legs does not result in any abnormal reaction, whereas stimulation of the skin of the arms immediately results in typical extensor tetani in the hindlegs.

These observations led me to the view that true strychnine tetani appear when both sensory and motor spinal mechanisms are strychninized.

de Boer has criticized these experiments, which were confirmed by Beritoff (2, p. 190), stating that he has been unable to corroborate them. Actually this observer confirmed my observations, in as much as he also produced typical strychnine spasms on simultaneous strychnin-

[1] S. de Boer (3a, 3b) has found that the reflex contractions on cutaneous stimulation after strychnine poisoning of the dorsal mechanisms electromyographically prove to be short tetani. It may be that the quasi-spontaneous contractions in the syndrome and even the few spontaneous ones still occurring after section and degeneration of the posterior roots are also of such a nature; de Boer has not investigated this point. Even if this were so, as probably may be the case, these very quick, small fascicular contractions are quite different from the genuine strychnine spasms.

ization of the dorsal and ventral spinal mechanisms in the frontlegs when the poison was applied locally to the dorsal and ventral surfaces of the arm region, in the hindlegs when the strychnine was applied locally on the leg region of the neuraxis (3b, p. 267 and 268). Only the above mentioned experiment with local poisoning of the dorsal mechanisms in the arm region and of the ventral mechanisms in the leg region of the frog was negative in the hands of de Boer. Apparently his animals were in rather poor general conditions after the operation, perhaps even before it; he confessed that he did his experiments on winterfrogs, which had not taken food for months. Whatever the cause for his negative result may be, it may be well to emphasize that a positive result can never be invalidated by a negative observation, unless it can be shown that the positive result is due to some fault in the method or technique used. As de Boer has not proved that such a factor was present in those particular experiments of mine, his negative result in this special instance need not, therefore, detain us any longer.

F. W. Fröhlich (10) has made the objection that tetani are absent in the syndrome of the dorsal mechanisms because the method of application used should result in only a partial, incomplete strychnine poisoning of these mechanisms. I must confess that I never understood the logic of this objection, as an argument to save the Verworn-Baglioni view, because I cannot see how the local application of strychnine to the ventral surface of the spinal cord can immediately augment the strychninization of the dorsal mechanisms. This, of course, must be the underlying reasoning if one wants to explain the prompt appearance of tetani in the experiments with combined dorsal and ventral strychninization on the basis of a specific action of strychnine on sensory mechanisms only. It is almost inconceivable that the minute quantity of strychnine applied ventrally should reach the dorsal gray matter almost instantaneously by diffusion through the soft membranes of the cord and nearly the whole width of the white and gray matter.

Fröhlich's objection is furthermore greatly invalidated by the following considerations: Such a local poisoning, even over only one-tenth of the length of the dorsal surface of *one* spinal segment is sufficient to bring forth the dorsal syndrome (without typical tetani) in the entire "strychnine segment zone" of that segment on both sides of the body. On the other hand the local strychninization of the dorsal surface of the whole lumbosacral cord does not produce tetani in the hindlegs.

It is clear from all this that in order to save the view taken by Houghton and Muirhead, Verworn and Baglioni we have to draw upon improb-

able, complicated reasoning, whereas the results obtained by these experimenters are easily explained on the simple assumption that the tetani in their experiments appeared because the application of the drug was not strictly local, but resulted in a spread of the poison. In the experiments of Houghton and Muirhead with application of the strychnine around the arm region it spread in all probability towards the ventral surface of the leg region and in the case of the application around the leg region it presumably spread towards the ventral surface of the arm region. On the basis of my experiments, it is clear that in the first instance tetani would appear in the legs on stimulation of the skin of the arms and that in the second case they would be observed in the arms on stimulation of the skin of the legs. Anyone familiar with application of watery fluids on the surfaces of the central nervous system, on reading Baglioni's description of his application technique (see above), will immediately suspect that in his experiments there is no real localization of the drug, but a spread of the strychnine solution all around the spinal cord. Then we have actually simultaneous strychninization of dorsal and ventral spinal mechanisms and true tetani will be observed.

We, therefore, still maintain that strychnine tetani occur and occur only when both the dorsal, sensory or better, connector spinal mechanisms and the ventral, motor spinal mechanisms are poisoned. There was, however, left one not quite satisfactory point in this view, namely, that strychninization confined to the ventral mechanisms, though a necessary co-condition in the production of tetani, did not yield appreciable symptoms; no sensory disturbances, no hyper-reflexia, no symptoms of motor excitation. As a matter of fact, in several of these experiments some fibrillar and fascicular contractions in various muscles were observed, but they were taken to be caused by the mechanical irritation and perhaps a slight damage of the cord and its ventral roots during the ventral application of the poison. Still, strychnine must act on the ventral mechanisms, since only its combined action on the ventral and dorsal mechanisms results in tetani.

In view of this not quite satisfactory state of evidence, I have recently repeated these experiments on the low spinal preparation of the cat and monkey (transection at Th. XIII or L. I) with a modified, improved technique of ventral application of the strychnine solution. In most of these experiments the animal was decerebrated after the spinal transection, to have a quiet preparation without anesthesia.

In the earlier experiments the poison was applied ventrally by lifting the cord with a broad spatula (after severance of the denticular liga-

ments) and painting the strychnine solution on the ventral surface in the usual way, or by pulling a small oblong piece of cotton wool soaked in the solution along the ventral surface of the cord. Obviously this method has its disadvantages, and in these recent experiments I followed another technique; viz., the denticular ligaments were also cut on both sides of the exposed lumbosacral cord, which makes slight displacements of the cord much more easy. Then minute drops of the usual strychnine solution (coloured with toluidine blue) were applied between the ventral surface of the cord and the underlying dura mater, one for each segment. This was done by the use of a tuberculin syringe and a very fine, blunted needle.

With this technique hardly any handling of the cord is necessary and in all experiments distinct symptoms of motor excitation occurred within a few minutes after this local ventral poisoning. They consist of quick, short, twitchlike, irregular contractions of various muscles, which may be strong enough to give rise to changes in position of the various parts of the hindlegs. Typical tetani and symptoms of hyper-reflexia were absent. When after an interval of from ten to fifteen minutes no augmentation of the motor symptoms was observed, the dorsal surface of the lumbosacral cord was locally strychninized, which immediately was followed by a strong hyper-reflexia and the appearance of typical tetanic convulsions of the hindlegs.

In all probability the negative results of the ventral poisoning in my earlier experiments were due to damage of the cord in handling it to obtain access to its ventral surface. Another factor may have been also that the strychninization in these new experiments was started long after the cessation of the anesthesia, which was stopped as soon as the animal was decerebrated.

The only question which remains for discussion is: which are these ventral and dorsal spinal mechanisms upon which strychnine acts?

In 1910 I thought that the nervecell bodies, the perikarya (Sherrington) are the seat of action of strychnine, and I am still of the same opinion. A. Fröhlich and H. H. Meyer (9) took the view that the connector neurons are responsible for the typical central strychnine effects and that strychnine tetani appear when not only the cell bodies of these neurons in the dorsal horns are poisoned (by dorsal application), but also their collaterals and teledendra in the ventral horns of the gray matter (by the ventral application of the poison). We are aware, that convincing evidence to decide between the two possible views is not available at present, and, let us add, will not very easily be obtained for

obvious reasons (the close, microscopical vicinity of the two structures). But there are some experimental facts at hand which seem to me of some importance in this respect:

It was stated above that the striking paresthetic disturbances of the dorsal syndrome still occur when the posterior roots of the poisoned segment and those of many adjacent segments have undergone Wallerian degeneration by section three weeks previous to the application of the poison. This fact necessarily leads to the conclusion that these symptoms cannot be brought about by action of the strychnine on the intraspinal portions of these root fibres and their terminal ramification or synaptic endings, but originate in the strychninization of the nerve cells of connector neurons situated in the dorsal gray matter. It is only reasonable to assume that these disturbances as well as the other symptoms of the syndrome, also under normal conditions, i.e., with intact posterior roots, are caused by the action of the strychnine on these nerve cell bodies. We know since the days of Magendie and Meyer that strychnine does not produce its typical effects by action upon peripheral nerve fibres. These facts together with the absence of any experimental proof that the intraspinal fibres and teledendra of the connector neurons should behave differently against strychnine than peripheral and intraspinal sensory and motor nerve fibres, leads me to prefer the first hypothesis rather than that of Fröhlich and Meyer, the more so as no reasonable and serious objections have been made known which would tend to invalidate it.

I am, therefore, still of the opinion that typical strychnine tetani occur only when the cell bodies (perikarya) in the dorsal horns of the spinal cord are poisoned simultaneously with the cell bodies of the motor cells in the ventral gray matter of the spinal cord.

BIBLIOGRAPHY

(1) BAGLIONI, S. Arch. f. (Anat. u.) Physiol., 1900, Supplem. vol., 193.
(2) BERITOFF, I. S. Folia Neurobiol., 1913, vii, 187.
(3) a. DE BOER, S. Arch. f. exper. Path. u. Pharm., 1923, xcvii, 30.
b. Ibid., 1924, civ, 265.
(4) a. BREMER, F. Arch. internat. de physiol., 1922, xix, 189.
b. Ibid., 1925, xxv, 131.
c. Mémoires de l'Académie royale de Médecine, Bruxelles, 1926, xxii, no. 8.
(5) a. BREMER, F. AND RYLANT, P. C. r. Soc. Biol., 1924, xci, 110.
b. Ibid., 1925, xcii, 199.
c. Ibid., 1925, xcii, 1329.
d. Ibid., 1925, xcii, 1331.
(6) CUSHNY, A. Quart. Journ. exper. Physiol., 1920, xii, 153.

(7) a. Dusser de Barenne, J. G. Folia Neurobiol., 1910, iv, 467.
b. Ibid., 1911, v, 112.
c. Ibid., 1911, v, 342.
d. Ibid., 1912, vi, 277.
e. Ibid., 1913, vii, 549.
f. Monatschr. f. Psych. und Neurol., 1913, xxxiv, 523.
g. Amer. Jour. of Physiol., 1929, xc, no. 2.
(8) Dusser de Barenne, J. G. and O. Sager. Zeitschr. f. d. gesammt. Neurol. und Psychiatrie, 1931, cxxxiii, 231.
(9) Fröhlich, A. and H. H. Meyer. Arch. f. exper. Path. u. Pharm., 1915, lxxix, 55.
(10) Fröhlich, F. W. Zeitschr. f. allg. Physiol., 1911, xii, 69.
(11) Hering, H. E. Pflüger's Arch., 1893, liv, 614.
(12) Houghton, E. M. and A. L. Muirhead. Med. News, 1895, lxvi, 612.
(13) Liddell, E. G. T. and C. S. Sherrington. Proc. Roy. Soc. London, 1925, B. xcvii, 267.
(14) Magnus, R. and C. G. L. Wolf. Pflüger's Arch., 1913, cxlix, 447.
(15) Meyer, H. Zeitschr. f. rationelle Med., 1846, v, 257.
(16) Miller, F. R. Nature, 1926, cxvii, 486.
(17) a. Poulsson, E. Arch. f. exper. Path. u. Pharm., 1890, xxvi, 22.
b. Hefter's Handbuch der Pharmakol., 1920, ii, 353.
(18) a. Sherrington, C. S. Phil. Trans. R. S. London 1898, cxc, 160.
b. Integrative action of the nervous system. 1906.
(19) Stilling, B. Untersuchungen über das Rückenmark, 1842, p. 46.
(20) Verworn, M. Arch. f. (Anat. u.) Physiol., 1900, 385.

THE METABOLISM AND DISTRIBUTION OF WATER IN BODY AND TISSUES

EDWARD F. ADOLPH

Department of Physiology, The University of Rochester School of Medicine and Dentistry

How and why water is required, distributed, and utilized in the animal organism constitutes a chapter in physiology that is both ancient and recent. The importance of the subject was first conceived by the "humorists" of Greek medicine; it was the first physiological problem to be investigated quantitatively—by the use of the balance in the hands of the Alexandrians Menon and Erasistratus, and later of Sanctorius; while the partition of water in the body is beginning only today to reveal its importance in the animal economy and in clinical symptoms. The physiology of water may form an introduction equally to the understanding of the organism as a whole, of organ physiology, and of cell physiology.

The present review of experimental results emphasizes the data upon normal man. Where these are not available, data from pathological man and from some mammals are referred to. It must be kept in mind that differences among species and individuals in the quantitative relationships of water are great enough to discourage broad generalizations based upon observations from diverse sources.

A number of pertinent investigations are summarized in previous reviews (297, 260, 196, 319, 274, 288, 235, 256, 96, 323). The comparative physiology of water forms another aspect that has been elsewhere dealt with (46, 238, 9, 280). The present account aims to define certain problems of general importance as they are illustrated in the mammalian body.

Water is the most abundant constituent of every tissue and is taken up and rejected more rapidly than any other substances. Why is it necessary to replace that which already abounds? Any marked deviation from the normal water content may have serious and various physiological consequences. A great deal is known regarding the parts taken by water in the metabolism of tissues. But the physical relationships of water in the tissues, which are probably of more consequence than the

chemical relationships, are less well understood. To regard water as merely the medium in which chemical changes are occurring is no longer adequate to the physiologist.

The physiology of water may be considered under two headings; the exchanges between body and environment, and the distribution among the tissues.

EXCHANGES OF WATER BY THE BODY. The supply of water to the organism may be described in terms of rates of intake and output through the several channels, and of the rôles of various influences upon the hydration of the body.

A. *Paths of Intake.* Water is normally all taken into the body by mouth, but much of it is disguised by mixture with other foods. Whenever ingestion becomes impossible it is very difficult to supply water fast enough. Rectal administration, gastroclysis, parenteral infusion and subcutaneous infusion are all inadequate for any purpose except temporary or experimental use. It is well known that water may be absorbed directly into the outer layers of the mammalian skin, but the passage of significant amounts inward through the skin does not occur.

1. *Alimentary ingestion.* The drinking of pure water accounts in most animals for only a minor part of the water taken by mouth. Solid foods and even "dry" foods contain very large proportions of water, and the feeling of moisture or dryness in the mouth is not a perfect criterion of the actual water content.

Water balances may be estimated in part from tables of food compositions. It must be recognized that particular foods vary greatly in water content according to their sources and to their modes of preparation. It is so far quite impossible to state with precision the water content of any one food, because analyses depend upon methods of desiccation (285, 231). The water contained in a food is present in several forms, ranging from pure solvent, through water of crystallization, of colloidal hydration, and of secondary valence, to water combined with carbon to form carbohydrate and the like. Until the conditions are known under which each form of water exists, analytical distinctions must remain arbitrary.

A great variety of foods and drinks depend for their physiological value wholly upon their content of water.

2. *Alimentary absorption.* Water once taken into the digestive tract does not become available to the tissues until many minutes later. In some circumstances it never becomes available, being either vomited or else passed through the alimentary tract and evacuated as diarrheic feces.

It is fairly certain that slight amounts of water may be absorbed within the stomach, but the conditions for this absorption are poorly defined. The amount so absorbed may be taken as negligible. Most absorption occurs in the upper small intestine, though the part played by the colon is equally requisite for normal alimentation.

The absorption of water in the intestine depends upon a variety of physico-chemical conditions. These have been studied upon operated dogs by introducing known solutions in measured amounts into intestinal loops (116, 229, 158, 248, 330a). It was found that water is ordinarily absorbed as a solution whose osmotic pressure is nearly equal to that of the blood. If the food solution be more concentrated than this, water comes from the blood into the lumen of the gut. If the fluid be dilute, absorption begins immediately, but at the same time various solutes pass from the blood plasma into the fluid. The nature and proportions of the solutes in the lumen are highly significant in controlling the rates of absorption. Whether energy is ever used locally in intestinal absorption has not been decided, since the forces of effective osmotic pressure of the blood have not been evaluated in the presence of the peculiar permeabilities of the intestinal wall.

Obviously water is absorbed most quickly when taken alone, or with food that after digestion forms a solution of less osmotic pressure than the blood plasma. The advantage to the body of taking water in amounts needed at the same time with food is therefore evident.

3. *Alimentary secretion.* Water is copiously secreted by the glands into the mouth, stomach and intestine. The quantities secreted are highly important, and have been variously estimated, totalling from 2 to 10 liters in 24 hours (table 1) (259, 260). The individual estimations depend chiefly upon operative animal experiments or upon studies of fistulous individuals, and such conditions are known to influence the rates of secretion. The quantities vary greatly with food intake and diverse bodily characteristics and activities. The rates of secretion become particularly important as modifiers of water balance in the several conditions in which the secretions are not reabsorbed. As long as the water remains in the alimentary tract the tissues are relatively dehydrated.

4. *Administration by clysis.* Introduction of water by paths other than alimentary usually requires use of a solution, and in man the use of aseptic precautions. Ordinarily such water will give most relief of dehydration when made suitable in concentration, not by salt, but by some combustible material such as glucose. If a salt is used, a hypo-

tonic solution of sodium chloride is preferable. If the body is in normal water balance, the fluid will usually be promptly excreted unless it contains isotonic sodium chloride (64, 72, 192, 193, 127, 246, 31). The best retention occurs when whole blood is furnished from another individual, even though it be introduced not into the blood but into the peritoneum or under the skin.

B. *Formation Within the Body.* Water is formed in two chief ways; by the polymerization or synthesis of many compounds (*i.e.*, the reverse of hydrolysis), and by the oxidation of any substance containing hydrogen. In addition, these chemical reactions, and some others as well, liberate certain amounts of water which were acting as solvent to, or were otherwise accompanying, these materials.

TABLE 1

Rates of turn-over of water by various organs in a 70-kilo human individual, in cubic centimeters per 24 hours

ORGANS	MINIMUM	LIBERAL
Salivary glands	500	1,500
Stomach	1,000	2,400
Intestinal wall	700	3,000
Pancreas	700	1,000
Liver (bile)	100	400
Lymph	700	1,500
Total recovered by body	3,700	9,800
Kidneys	600	2,000
Colon (feces)	50	200
Skin (insensible loss)	350	700
Sweat glands	50	4,000
Mammary glands	0	900
Total lost from body	1,050	7,800
Total turn-over	4,750	17,600

1. *Hydrolysis.* Most of the changes in the alimentary tract are such as to absorb water, since the products of digestion are of lower molecular weights than the original food materials. The amounts of water thus absorbed can be calculated stoichiometrically from the amounts of each type of food in the diet. They constitute 5 to 20 per cent of the dry weight of the food substance.

A large part of this water is liberated again after the translocation of

food from the digestive lumen to the various tissues where the material is dehydrolysed. In general the permanent forms in which materials exist in the body, whether structural, fuel, or storage, are like the ingested forms before they have been digested. Similar hydrolyses may occur in the tissues, and while the water formed and the water absorbed balance one another in the whole organism and over a long period of time, the individual tissue balances vary considerably at successive times before and after absorption. Direct data upon this subject do not exist.

2. *Oxidation.* Water formed by the oxidation of hydrogen-containing compounds is often termed "metabolic water." It can be calculated for any long period of time after ingestion as the "potential water of foods." In the case of nitrogenous foods, a correction must be made for the fact that a small proportion of hydrogen is protected from oxidation by the formation of urea. It happens that oxidation produces water nearly in proportion to the calorific value of the food, yielding 10 to 14 grams of water per 100 Calories (see table 3).

Since water is formed oxidatively within cells, it is possible that temporarily this water may have special significance at the site of its formation. Undoubtedly this water sets up a minute permanent gradient of concentration, causing a predominance of osmosis from the cell to the pericellular fluids.

3. *Storage and liberation.* While fat is stored without significant amounts of water, certain lipemic conditions are accompanied by water retention (93). But in contrast to other adequate diets, one high in fat may lessen the water content of the body (42, 36). Carbohydrate and protein are said to be deposited with definite amounts of water at the site. The amounts deposited have been measured only roughly, by feeding equivalent amounts or supplements of carbohydrate or of protein while measuring the retention of water and nitrogen (42, 234). For carbohydrate the only estimate is the one of Zuntz (340) that 3 grams of water are held with each gram of glycogen. This estimate is founded on very poor data (52). For protein the data are less definite. Sudden changes to a protein diet from a mixed one give losses of water from the body (42). The ratio of protein to water in each tissue is fairly constant, and if it be assumed that protein is deposited uniformly among the tissues of the body, then roughly 3 grams of water are deposited with each gram of protein, or 19 grams with each gram of nitrogen.

The relation of water to food is also manifested during catabolism. The normal rate of water loss depends upon the liberation of water from the tissues rather than upon the demands of excretory organs (7).

Hence the ratio of water loss to oxygen, carbon, or nitrogen metabolism during post-absorptive periods is a direct measure of storage requirements.

During the starvation almost no water is required by the body. Thus, dogs ingested only one-fourth the usual amount of water upon interspersed days when no food was given (157). The fact that less water is excreted in urine at night than in the daytime is largely dependent upon the diminished catabolism at night.

C. *Paths of Output.* While there is only one important path of intake of water, there are many paths of output. Within certain limits these excretory organs may compensate for one another.

1. *Urinary.* This path responds the most readily to the requirements of the body. But it is impossible for the urinary rates of water excretion to decrease much below the usual rates, for when the body is short of water the kidneys still take the amounts required for the excretion of dissolved materials. The amounts of water used for this purpose appear to depend upon the fact that the kidneys cannot concentrate these materials beyond a certain osmolar limit.

The changes in normal rates of urinary output with age have been summarized for man (126). It appears that the rates are correlated most accurately with body surface, and that the average output is 1500 cc. per square meter per day. Since the rates of other catabolisms are also proportional to body surface, this correlation was to have been expected. The minimum output per day on a dry diet is of the order of 400 cc. of urine per square meter, showing that *ad libitum* intake of water is on the average at least one liter above the most absolute needs.

2. *Pulmonary.* The rate of water loss from the lungs and nasopharyngeal passages depends upon the state of the inspired air. Expired air is ordinarily saturated with moisture at about 34°C. (332, 187). The air therefore carries away as much water as is required to saturate it. If the inspired air should have a dew point higher than the deep temperature of the body, then water would condense in the respiratory passages and be absorbed.

Under basal conditions in adults the pulmonary loss averages 36 per cent of the extrarenal losses, or about 300 cc. in 24 hours (35). The loss increases measurably with such activities as reading aloud and singing (265). The loss is speeded up in any condition, such as exercise or other acidosis, where the respiratory minute volume or ventilation rate is increased (332, 109, 289, 50).

3. *Sweat.* Under basal conditions the sweat glands are generally believed to be inactive, so that loss of water in sweat may be considered separately from "insensible cutaneous loss" in the restricted sense. These glands are primarily for emergencies, and take water from the body in order to dissipate heat by evaporating it. In exercise, exposure to high temperatures, and fevers, the losses of water as sweat may be very great. The excess heat produced in exercise is lost up to the extent of 50 per cent or more by the evaporation of sweat (fig. 1) (186, 138), the heat lost in evaporation being calculated from the fact that for

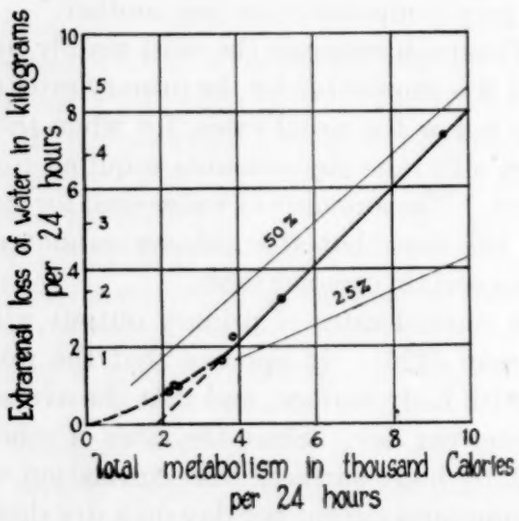

Fig. 1. Losses of water by evaporation in relation to total heat production. The data are taken from Loewy (186). The supplementary scale of ordinates represents thousand Calories of heat per 24 hours dissipated by evaporation. The slope of the straight portion of the curve indicates that 55 per cent of the energy used above the basal is dissipated by evaporation. The two light lines indicate the relation between evaporation and heat production if 50 per cent and 25 per cent of the total heat were dissipated by evaporation, respectively.

every Calorie so lost, at least 1.73 cc. of water is secreted. In hot atmospheres the amount of sweat is approximately proportional to the effective temperature above 32°C. (8). In fever there is no simple rule to predict the water required for sweating, but it is closely related to the mean body temperature (134).

Most mammals other than man have few or no sweat glands, and use more exclusively other surfaces, especially respiratory, to dissipate heat by evaporation. In man the glands are congenitally absent in certain rare individuals.

4. *Insensible cutaneous.* The gradual and continuous loss of water from the skin has been recognized for centuries, but the question is not yet fully decided whether the sweat glands have anything to do (151, 219) with that loss which goes on at ordinary temperatures. In man under basal conditions 200 to 400 cc. per square meter per day are thus lost (293, 37, 137, 150). The loss is equally or nearly as great in persons without sweat glands (188, 251), but in them the skin losses can be increased only slightly by muscular exercise or by temperature. Investigations have been peculiarly deficient in showing by direct measurement that the loss is faster in dry atmospheres of the same effective temperatures, but this must be the case from the fact that dry atmospheres have more cooling power (143). While a minimal amount of insensible loss may be necessitated by some active process, there is no evidence that prevention of loss by varnishing the skin is deleterious.

Insensible losses are often measured by the differences of all the other gains and losses over 24 hour periods. But they are best determined over short periods with a Lombard (189) or Sauter (37) balance. The moisture of the expired air must be collected separately, absorbed and weighed; and the entire body surface should be exposed to an atmosphere of known temperature, humidity and rate of air motion. Methods have also been devised for measuring the speed with which air in contact with a limited area of skin takes up moisture (331, 126 p. 52), but here too the air should be of known condition and continually renewed at the body surface with a known velocity. It has been demonstrated that evaporation is more rapid from some areas of skin, as the arms and legs, than from equal areas in other regions, as the trunk (185).

Insensible losses are markedly diminished by limitation of water intake or other dehydration (298, 282, 51, 77, 194) and increased by water drinking (75, 220, 51). Psychic influences and reflexes are important factors (220), probably working in part through the circulatory system. The temperature of the skin must always be considered a primary regulating influence (185, 8).

Studies of insensible evaporation from the skin have been made on a few mammals other than man. Dogs (264) and rabbits (153, 205, 198) yield results parallel in many respects to those obtained on man.

5. *Alimentary.* Water is actively secreted into the colon independently of digestive functions, accompanying substances such as calcium and sulfate that are largely excreted by this channel. In some forms of diarrhea it increases enormously in amount. Normal human feces contain about 80 per cent of water, amounting to about 100 cc. per square

meter per day on an average mixed diet. But most of this normal water remained with the food substances in the digestive tract with the unabsorbed food residues.

6. *Special paths.* Significant amounts of water may be lost from the body through tears, nasal secretions, spitting and drooling, hemorrhage, breaking of skin vesicles or blisters, and secretions of the genital tract. In operative conditions evaporation and drainage of fluid from subcutaneous and serous reservoirs may involve large masses of water, which cannot be ignored when a water balance is being reckoned or a water requirement calculated.

7. *Partition of outputs.* Every influence that affects the output of water by the body creates a shift in the relative and absolute amounts of loss through each channel. In man the partition of losses may be represented as in table 2. There is no conclusive evidence that the percentage of the total loss that goes out through any one path changes significantly with age or body surface.

D. *Turn-over Rates.* The estimation of total water balance over a period of a few hours is not difficult and yields considerable information. The tabulation and calculation of the various intakes and outputs, and of the retention, has been presented in a number of papers (213, 232, 334). One of the things learned from such an estimate is the huge amount of water required to keep up the body's supply. Most of the ordinary losses from the body represent water that comes to be eliminated because the other materials with which it was associated have been catabolised. But when water intake is limited, then water is still appropriated by the various excretory channels at rates which are relatively little reduced from the normal.

Water stands far above every other substance in rate of turn-over by the body, whether it is compared absolutely in grams of molecules per day, or relatively in per cent of the body's content (9).

1. *Standard requirements.* It is possible from the relationships that have been described above to estimate with fair accuracy how much water is required by the human body or by another kind of mammal. In this requirement factors must be considered for size, activity, temperature influences, and other unusual circumstances (table 2). Fortunately, water requirements are closely related to energy, nitrogen, and salt requirements, and are chiefly influenced by the same factors.

Urinary output of water is necessitated by the solutes that demand excretion. These are excreted in such a manner as to form normally a total urinary concentration of about 0.6 osmolar in man (160, 56, 183,

182, 61, 74, 7) and a considerably greater concentration in most other mammals (13, 61, 68). The voluntary intake of water has been studied as a function of solute excretion in rats (110). In extremely dehydrated conditions concentrations of 1.2 to 1.4 osmolar are attained in man.

Pulmonary output is calculated from the saturation capacity of the inspired air and from the temperature of the body. Sweat is related in amount to the energy expenditure for exercise and to atmospheric conditions. Insensible cutaneous water losses constitute 24 to 30 per cent of the energy production under the most usual circumstances (35, 293, 37, 78, 205). Alimentary outputs are nearly always small, and so far can

TABLE 2

Coefficients for calculating the water requirement of an individual, in cubic centimeters per 24 hours, from the surface area of the body (A) and the energy requirement (E)

A is measured in square meters, E in large Calories.

UTILIZED FOR	MINIMUM	AVERAGE TO LIBERAL
Growth (or storage)	0	15 A to 30 A
Urinary	400 A	1000 A to 1500 A
Fecal	30 A	90 A to 150 A
Basal extrarenal	250 A	390 A or 1.73 × 0.25 E basal
Exercise (sweat)	1.73 × 0.4 E excess	1.73 × 0.55 E excess
Temperature control (sweat)	0	1800 A (atmospheric °C–32°C.)
Total (approximate)	2100	3400 to 5000

be related only to the roughage content of the diet. Special modes of loss are usually measurable. Thus all the chief factors can be taken into account, in some such way as is shown in table 2.

An average man requires, therefore, something like 3400 cc. of water per square meter per 24 hours, making allowance for three times the basal consumption of energy. A typical 3000 Calorie diet as consumed at the table contains perhaps 2000 cc. of preformed plus potential water. Hence 1400 cc. remains to be drunk as such. Roughly a convenient liberal standard of total water intake for any mammal is 1 cc. per Calorie.

2. *Factors influencing requirements.* The influence of food upon

water requirement can best be outlined by means of table 3. Each food is characterized by a content of preformed water and a calculable amount of potential water. But the heat produced by each Calorie of food energy requires 24 to 30 per cent dissipation by evaporation, and the metabolic residues, chiefly ash and urea, require water for their excretion. The principles used in making these calculations are substantiated by measurements of voluntary water intake in rats (253, 110) and mice (40). But the exact values derived have not been checked experimentally; this can only be done on man, and the amounts of water to be traced are relatively small so that the accuracy of measurements must

TABLE 3

Grams of water needed for the complete metabolism of 100 Calories of some food substances

One Calorie of protein requires 3.0 cc. of water for the excretion of the urea and sulfate formed from it. One gram of ash requires 65 cc. of water for its excretion.

FOOD MATERIAL	PREFORMED WATER	GAINED BY OXIDATION	LOST IN DISSIPATING HEAT	LOST IN EXCRETING END-PRODUCTS	DEFICIT OF WATER
Protein	0	10.3	60	300	350
Starch	0	13.9	60	0	46
Fat	0	11.9	60	0	48
Beef, sirloin	25	11.3	60	119	143
Fish, cod	120	10.4	60	382	312
Eggs, hen	47	11.1	60	154	156
Milk, whole	127	12.5	60	123	43
Bread, white	14	13.2	60	69	102
Apples, fresh	150	13.9	60	56	−48

inevitably be poor compared to the accuracy of calculations. Complications arise with the initiation of sweating, for considerable amounts of ash and urea are contained in sweat, relieving the kidneys of the necessity of excreting them.

Fortunately, too much water ingestion is rarely deleterious to the organism, and so it is usually enough in practice to insure merely that the minimum requirement of water is available to the organism.

3. *Prenatal water relations.* Exchanges of water before birth are entirely between mother and fetus. There are no data to indicate whether the percentage water content of the fetus changes from hour to hour with that of the mother, but such is the customary assumption.

In human mothers the retention of water increased markedly during pregnancy, particularly during the last half (338, 339). The generalized tendency to edema during pregnancy may be the cause or the effect of this.

Nothing is known about the capacity of the placenta for exchanging water. The only data bearing upon the interchanges that occur there are the concentrations of fetal blood and maternal blood simultaneously sampled. At the time of birth the total osmotic pressure of fetal blood is about 5 per cent higher than that of maternal blood (320). The plasma protein concentration is considerably lower (210).

The fetus may be regarded as a fast-growing tissue of the mother's body. The other placental tissues also enlarge rapidly and so lead to marked retentions of fluid. The turn-over of water may be correspondingly large, and thus the total water requirement of the mother may become disproportionate to her surface area.

Analyses of aborted human fetuses (104, 53, 216, 59, 309) and analyses of surgically obtained fetuses of other animals (104, 227) indicate marked progressive reduction of relative water content as the fetus grows.

E. *Water Content.* While the turn-over of water is the tangible quantity, what really matters to the organism is its content of water.

The relative content of water decreases slowly with age both before and after birth. Data have been recorded for the whole bodies of men (41), dogs (113, 305, 120), cats (305), rabbits (104), cattle (227), swine (227), guinea pigs (227), rats (133, 227), mice (227), and a few other species. The data are most consistent when referred to the weight of fat-free tissue. The change that characterises the fetus continues during the infantile period; the change gradually slows but extends even into senility. Undoubtedly the lower percentage water content of the adult is related to the greater rigidity and tensile strength of mature tissues.

1. *Influences upon water balance.* Retention or loss of water has been reported to occur in response to a number of experimental and normal factors. Two criteria have been used: analyses of the entire body to give relative water contents, and changes of weight of individual animals under controlled conditions. Some of the factors studied are: pyloric obstruction (83), starvation (34, 321, 322), low-protein diet (36, 279, 76, 159, 107, 286), high NaCl intake (299, 333, 105, 44, 241, 214, 318, 6), depletion of plasma proteins (178, 179), life at high altitudes (123, 65, 300), administration of anterior pituitary extracts (82), deficiency of

vitamine B (257), menstruation (135, 268), and pregnancy (338, 268, 29). Every one of these factors has also been identified as modifying the concentration of the blood, and some are correlated with changes of concentration of particular tissues and of the volume of the blood. While the concentration and volume of the blood furnish the immediate stimulus to retention or dehydration of all regions of the body in some instances, the factors of blood pressure changes, hormones, and nervous mechanisms nearly always come into play. Hormonal coördinations are best illustrated by the effects of intestinal extracts (66, 14), thyroxin (101, 272, 108, 131, 118), insulin (284, 96), liver extracts (217, 3), and brain extracts (48). The rôle of the central nervous system has long been appreciated in relation to abnormal loss of water through the kidneys (38, 91, 4, 47, 54, 252). It is possible that the autonomic nervous system also is of first importance in water regulation (242), though it is not clear how much of its influence is purely vascular.

The essential relations of the body to water intake can be described very briefly. If water is drunk, an equal amount is excreted usually within two or three hours in the urine, in addition to the water excreted at the normal rate. Meanwhile the insensible loss of water increases in rate so that the body often ends up by having as little or less water than it would have had if none had been ingested. The same result is obtained from drinking a solution of almost any salt in a concentration that is isosmolar with the blood. But if the salt is sodium chloride or sodium bicarbonate, then little or no diuresis results and the excess fluid is eliminated very slowly (during 24 or 48 hours) (333, 55, 19). If protein, carbohydrate, or fat is taken with the water, certain amounts of the fluid are retained and stored with the food until the food becomes oxidized (234, 222, 223, 215, 206, 240). If the water is taken before food it will be promptly excreted; whereas taken with or soon after the food or salt it may be retained. If water alone is taken, particularly after excessive loss of salts has occurred, such as by sweating, the dilution of the tissues may be sufficient to bring on intoxication or cramps (226, 261, 316, 262, 124).

For many normal persons, such as tend to be obese or enuretic, and for persons with certain diseases, such as eczema (212) and idiopathic epilepsy (208), it is desirable to drink under conditions such that any water in excess of the most pressing needs will not be retained.

2. *Effects of changes in water content.* Dehydration resulting from low water intake or other processes produces many effects that are quite obvious. The sense of thirst normally detects dehydration first, but is

readily interfered with by experimental or pathological conditions. The superficial signs of lack of water are dryness of the skin, shrinkage of subcutaneous tissues, decreased deformability of skin, and decreased rate of urine excretion. Diminution of blood volume (68a, 154, 71, 207) and shrinkage of other tissues (162) follow. The oxygen consumption of cats is increased during the first few days of thirst (58).

A diminished water supply leads to some economy of water in insensible loss as well as in urine. But sweating, diuresis, and certain tissue demands are met with very little saving by the dehydrated body.

The changes of hydration of the body as a whole are known to be influenced by a number of bodily activities such as muscular exercise, sleep, and exposure to extreme temperatures. These are additional to the activities, to be mentioned below, that are identified closely with particular tissues.

The immediately increased availability of water in the body is known, on the other hand, to affect many functions, such as rate of renal blood flow (167), rate of total blood flow (149, 124a, 117), rate of oxygen consumption in man (36, 149, 190a, 124a, 59a) but not in dog (266, 191), rate of nitrogen elimination in man (1, 335, 145, 197) but not always in dog (221, 298, 136, 191, 145, 197), rate of growth in weight (146), and rates of salivary (72, 67), gastric (239, 301) and bile (292) secretion.

3. *Tests of hydration.* All measurements of the state of the body with respect to water must be relatively empirical until the mechanisms of regulation are understood.

The most obvious test of adequate water supply is to give a large quantity of water by mouth and to see whether any of it is retained, or at least whether the urine becomes dilute. In most cases this indicates whether or not the body is "saturated." Body weight changes may be used to detect water retention.

The condition of the skin may be measured by several means, such as pitting (273), elastometry (114), standard wheal formation (90, 111, 243), and intradermal salt or serum absorption (12, 250).

Intraocular tension (86a), dryness of the mouth (297), rate of reflex salivary secretion (67), and other measurable quantities are correlated with the state of the body with respect to water.

Measurements of blood concentration and blood volume are applicable. The various methods that have been used are illustrated by the following determinations in the study of water drinking on man and dog: erythrocyte count (70, 156), hemoglobin concentration (249, 199, 201, 316, 124), hematocrit (84), dry content (43, 255, 245), refractom-

etry of serum (99, 317, 249, 70, 156, 124, 30), viscosity of serum (70, 124), freezing point depression (249), evaporation rate (195), electrical conductivity (336, 245, 30), chloride concentration (249, 245, 316), specific gravity of plasma (221), and volume of plasma (177, 124, 201, 84). In the few instances reported where no blood dilution has been found after drinking, it can be supposed that the water ingested was too small in amount or too slowly absorbed.

All methods indicate something of value, but most of them can be accurately interpreted only when the same individual is being studied from day to day or from minute to minute. Upon the patient or the experimental animal that has never been controlled before, the only adequate test of water balance is one in which the mechanisms of water regulation receive a graded stimulus, such as the ingestion of a large volume of water, and then the response is measured in rate of urine excretion (323 p. 164) or time curve of blood dilution (255). Other tests are little superior to gross observations of physical appearance.

The interrelations between the net water content of the body, the rate of turn-over of water, and the rates of intake and output of water by each channel, obviously offer varied possibilities for regulation. The control of these paths can in its nature only be complicated. But the conclusion must be kept in mind that the water balance of the body depends upon not only the work of the organs of intake and output but also the relationships of each tissue within the body.

WATER IN TISSUES. A quantitative knowledge of intermediary water exchanges becomes imperative when it is realized that water enters into most of the regulatory activities of the tissues and organs. How water is engaged in vital processes must contribute to the elucidation of the physical bases of many physiological mechanisms. Upon the distribution of water in the body depends the translocation of other materials; to a considerable extent this relationship is reciprocal, and hence with the movement and content of water in each tissue are correlated many activities of that tissue. It is unfortunate that the only tissue whose changes of water content have been studied extensively, the blood, is perhaps the most passive of tissues.

A. *Physico-chemical Properties.* It is widely understood that water has a peculiar combination of properties that enhance its biological possibilities (139). Its high surface tension, its great tendency to form hydrates, its high dielectric constant, its high specific and latent heats, and its particular absorption spectrum for radiations are illustrative of what characteristics make it physiologically useful. Many of the

unexpected peculiarities of water are possibly accounted for by the fact that its molecular form is not ordinarily H_2O, but a mixture of H_4O_2, H_6O_3, etc. It is possible (15) to extend this fact to furnish hypothetical explanations of the colloidal hydrations, permeabilities, and phase maintenances of living cells. But thus far such explanations are not clearly implied by any experimental data. The most bulky phase in every cell is water, serving to promote the diffusion of soluble materials, but at the same time keeping its own molecules continually in motion. Even though the water content of the cell remain constant, the actual molecules that are present differ every hour. When the water content changes, as it is usually doing, more molecules come than go.

B. *Tissue Interchanges.* Many studies have been made on excised tissues, by measuring the rates of osmosis into them and by establishing conditions of aqueous equilibrium. The surfaces of exchange after isolation are very small compared with those *in vivo* where capillary walls bring together blood and tissues. Thus in a dog's gastrocnemius the capillaries have surfaces 200 times the external area of the muscle (161a). But from the rates of osmosis into isolated tissues, osmotic constants, which are proportional to the permeabilities of the tissues, can be calculated (10).

1. *Osmotic pressures.* When equilibrium conditions are sought it is found that an excised tissue is not in equilibrium with blood, but takes up water from it. Addition of salt to the blood or other immersing solution does little to prevent the endosmosis, but the addition of protein does. Rates of osmosis into dermis, umbilical cord, tendon (277), muscle (10), brain tissues (128) and blood (295, 296, 174, 175, 176) have been studied. The protein is efficient in opposing endosmosis because it is the chief or only substance present that does not penetrate tissue and so is able to exert its full osmotic pressure. It is nevertheless true that other constituents are able to exert part of their osmotic pressures, and to exert them over short periods of time during which they are diffusing toward new equilibria.

The lack of equilibrium between blood and tissue is accounted for by the presence while in the body of other forces or pressures in the tissues. These may be turgor pressures of single cells or elastic tensions of the whole tissue, which may change or disappear as soon as the tissue is excised.

The total osmotic pressures of various excised tissues have been measured by numerous procedures (129, 46, 141), particularly by determinations of freezing points and evaporation rates. But since these

osmotic pressures are not exerted across membranes having the peculiar and particular permeabilities of living ones, *effective* osmotic pressures are the important values to be desired. Their measurement depends upon the imitation of the *in vivo* permeabilities by foreign membranes, such as graded collodion or cellophane ones. Approximately the osmotic pressures exerted across the capillaries are those due to colloids, as Starling (295) recognized. What is thus measured at equilibrium probably comes fairly close to the effective value that exerts itself across a living normal membrane in a steady state. But where sudden changes of non-colloid osmotic pressure are produced, as in the intravenous administration of solutions, temporary changes of effective osmotic pressure may differ greatly from the measured ones. Thus, 5 to 25 per cent NaCl solution is effective in draining water into the blood of a dog (225, 328, 155, 11) because the salt cannot leave the blood as rapidly as the water can come into the blood, and because cellular impermeability either to cations or to anions causes inequalities in total ion concentrations that are dynamically effective until ions and water have both had time to redistribute themselves.

The colloid osmotic pressures of blood (295, 161a, 121, 269, 276), of lymph (184), and of serous fluids (276), have been measured and correlated with the concentrations and kinds of proteins present (269, 122). The colloid osmotic pressures of solid tissues can also be measured (86), and it is found that the pressures exerted by muscle are higher than those of blood plasma and of lymph, as might be assumed to be the case from the studies of osmosis in excised tissues.

2. *Permeability.* It was once usual to regard osmotic pressure as the chief or only factor concerned in the distribution of water among tissues. It is not the only factor, because it has been shown that hydrostatic, electrostatic, chemical, and deformational forces are present to sum algebraically with the osmotic pressure. Osmotic pressure is not always even an important factor, because few of the solutes are able to make their osmotic pressure effective so long as the living membranes are permeable to most of the solutes present. The resistance of a membrane or a tissue to the penetration of water, which is the reciprocal of its specific permeability to water or the rate at which water can pass through the tissue, probably differs greatly among tissues and in the same tissue at different times.

It is generally believed that the permeability of tissues to water does not change significantly with functional activities, but that the permeability to substances exerting osmotic pressure, particularly to proteins,

changes, and by this means the movement of water is modified. Direct evidence for such changes exists (168), but is scanty.

3. *Hydrostatic pressures.* It has been shown *in vivo* that the mean hydrostatic pressure present in the blood capillaries is approximately that required to compensate for the colloid osmotic pressure of the blood; hence this hydrostatic pressure prevents fluid from coming into the blood from outside the capillaries (60, 168, 170, 171). It must be equally recognized that the effective osmotic pressure of tissue cells would draw water in from the surrounding tissue fluid that is in dynamic equilibrium with the blood plasma, if there were no turgor within them. The capillary equilibrium is probably dynamic only in the respect that the various parts of the same capillary and near capillaries average their hydrostatic pressure, for lymph formation in the basal condition is very small, and in skeletal areas almost negligible. In fact, the production of water by metabolism in tissue cells just about supplies the amount leaving the tissue in the lymph stream. Of course, in departures from basal conditions many changes in these quantities occur. The rate of blood flow is believed to have no influence upon water exchanges so long as the oxygen tension is unchanged, because the pressure relations are not modified.

Thus it is evident that four *chief* quantities are concerned in the equilibrium between every cell and the tissue fluid about it, and between the blood and the same tissue fluid, or in the gross between the tissue and the blood plasma. These are, the blood's hydrostatic pressure within the exchanging vessels, the turgor pressure and deformational tension of the tissue around the capillaries, the effective osmotic pressure of the blood plasma, and the effective osmotic pressure of the tissue cells. All of these quantities, with the exception of intracellular turgor pressure, have been measured, though only approximately and usually as average values. Their balance is indicated in figure 2. It is believed by Landis (170) that the mean pressure of the blood in capillaries is equal to the effective (colloid) osmotic pressure of the blood plasma which circulates through them. If this is true, as it seems to be, it follows that the hydrostatic (turgor) and osmotic pressures exactly compensate one another both in the tissue fluid and in the tissue cells. If it is not true, then having a higher osmotic than hydrostatic pressure exerted, as Krogh (161a) believes, would necessitate that both the tissue fluid and the cells have the same absolute excess of osmotic pressure over turgor pressure. It cannot be true for any long period of time that the hydrostatic pressure is higher than the effective osmotic pressure in any one

of them. Moreover, the turgor pressure of the cell or of the capillary cannot fall below that of the tissue fluid without collapsing. It is evident, therefore, that the probable equality of hydrostatic and osmotic pressures within the blood capillary is not a proof that equilibrium of water exchange exists, but demonstrates that in a mean steady state the two particular forces happen to compensate one another. In the intestine a negative pressure approximately equal to the colloid osmotic pressure of the lymph or of the blood can prevent absorption of water (330a).

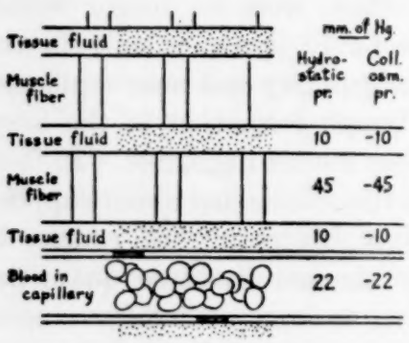

Fig. 2. Diagram representing the balance of forces concerned in the osmotic steady state in average mammalian muscle tissues. The approximate values indicated are taken from Landis (170, 171), Drinker (184), and Duff (86). The values for the colloid osmotic pressures may not correspond with the effective osmotic pressures in all instances; the value for the hydrostatic (turgor) pressure within the muscle fiber is wholly assumed.

In addition to these quantities, it is likely that two other forces have relatively small influences. Forces of chemical hydration and of electrostatic pressure can be demonstrated *in vitro* to exist, yet cannot be put into the algebraic sum of forces controlling water exchange.

4. *Chemical hydration* is credited to a great many of the substances found in living cells. Thus, KCl shows hydration of its dissolved ions; all carbohydrates whose osmotic pressures have been measured show hydration of molecules; the existence of a protein molecule that does not influence the water in which it is dissolved or suspended is almost inconceivable. In fact, it is hard to find any sort of compound that behaves like an "ideal" solute in water. On the other hand, whether it is correct to interpret all departures from the ideal as evidences of chemical hydration is doubtful.

Often the water of tissues that is held by forces of chemical hydration and in non-aqueous phases is termed "bound" water, though that term is also loosely used in other connections. Bound water has been empirically measured by various methods, each method carrying its own definition of what bound water is. Rubner's calorimetric method (267, 304, 254), Newton and Gortner's solubility method (233), Ostwald's viscosity method (164, 165), and Svedberg's (302) and Moran's (224) dilatometer methods are all interesting procedures, but depend

for their significance on how the results are calculated (125, 224). At the present time there is no good evidence that water is "bound" in significant amounts within any animal tissue.

5. *Electrical potentials* are found in any system containing local differences of ionic concentration. The potential differences can be readily measured, but they represent only one factor in the capacity of the system to support and maintain unequal distributions of all the various substances, including water. One sort of these is the calculable difference of osmotic pressure and of electrical potential directly due to a Donnan equilibrium. Experimentally the effect of electrical forces in transporting water can be measured by imposing known currents across a particular tissue (228). But in no specific instance have the electrical forces in a living tissue been evaluated in terms of equivalent hydrostatic or osmotic pressures.

C. *Redistributions of Tissue Fluid.* The study of the distribution of water within the living body is greatly influenced by the methods employed. No description of water balance is complete without measurements of the translocations of water among the tissues. Thus far it has been necessary to excise samples of the tissues to determine their relative water contents. The conditions of sampling are extremely important, for it has been found (11) that shifts of water are surprisingly rapid so long as the blood is circulating. Ordinarily the tissues slightly and progressively decrease in water content during the failure of the circulation. It must also be recalled that any sample of tissue includes blood, tissue fluid, and lymph as well as tissue cells. It is possible to estimate approximately the amount of blood included in the sample by measuring the quantity of hemoglobin present, since the amount of tissue hemoglobin even in red muscles is small and constant compared to that of the blood coursing through the muscle. The amounts of tissue fluid and of lymph contained in a tissue sample have not yet been estimated.

1. *Systemic factors.* The effect of cold upon water redistribution illustrates the rôle of vascular mechanisms. It was shown (20, 25) that the whole blood, not merely the blood in cooled capillaries, becomes anhydremic. Upon unilateral exposure of an animal to cold, the water that leaves the blood was found by analysis in the skin and muscle of the cooled area (130). The reason that the entire blood loses its water to a highly localized region was found in the fact that the arterioles strongly constrict, and subsequently the capillaries peripheral to them dilate. Perhaps the capillaries become more permeable to proteins as a result of the relative anoxemia (22), or else the capillary blood pressure increases

(171); in any case transudation of water is promoted in these particular areas of superficial tissue.

It has been known for a long time that rapid changes of blood concentration occur during many physiological activities. These have been taken to indicate shifts of water between blood and various tissues. This interpretation became more complicated with the realization that various organs such as the liver (297, 166, 202), spleen (27), and bone marrow (144), may serve special rôles in changing the concentration of the blood, and particularly that the blood in capillaries of the skin, spleen, and other organs may stagnate to a huge degree and so get out of equilibrium with the mixed blood. Hence it has become more desirable to fall back upon measurements of circulating blood volume, as the true indices to water exchanges by the blood. By one or both of these criteria it is known that a great many influences cause water to shift into and out of the blood. A number of these factors are listed in table 4. In addition, the concentration of the blood varies with species, age, and sex.

2. *Particular tissues.* Changes of the amount of water contained in the blood have been measured in relation to a number of individual organs and tissues. During continued contractions of a muscle within the body much water is removed into it (28). Some of this water forms tissue fluid and lymph, but some of it may enter the muscle fibers. It is possible that this translocation of water is chiefly due to the well-known increase of osmotic pressure in rapidly metabolizing muscle. But it may equally well be credited to increased capillary blood pressure or to changes of chemical hydration.

Most glandular tissues swell during activity (26, 28), and it is usually believed that part of this represents an increased water content of the cells. In addition the glands take water from the blood about as fast as they secrete it (28). The connection between the increased content of water and the increased production of secreted fluid and of lymph has not been traced in detail. All glands appear to reduce the rates of their secretion when the concentration of water in the blood plasma is diminished, and the reverse holds for at least the salivary glands.

The sensation of thirst, the diminished motility of the intestine, and changes in the properties of the skin, all of which accompany dehydration, may be taken as indicative of shifts of water between blood and other tissues.

By sampling tissues from anesthetized animals and analysing these tissues for their content of water, large and rapid changes of water are

TABLE 4

Some factors influencing tissue water distributions

PRIMARY FACTORS	SECONDARY FACTORS	SOME MODIFIERS	SOME EXPERIMENTAL PROCEDURES	CHANGES OF BLOOD CONCENTRATION MEASURED	CHANGES OF TISSUE CONCENTRATION OR VOLUME MEASURED
Tissue and turgor tensions	Tissue activities Contractures	Specific ions Hormones	Gravity Squeezing Suction Thyroxin Histamine Inflammation	Thompson (307) Dautrebande (73) Fujimaki (108) Dale (69) Underhill (313)	Waterfield (326) Mende (211) Ebbecke (89) Eppinger (101) Lewis (181) Underhill (315)
Capillary blood pressure	Blood volume Vasomotor system Venomotor system Capillary responses	Cardiac activities Autonomic activities Local temperatures Hormones	Exercise Vasomotor stimulation Surgical shock Adrenalin Pituitrin Inflammation Temperature	Böhme (45), Dill (79) Sherrington (287) Gasser (112) Ederer (92) Poulsson (244) Landis (172) Barbour (25), Marx (200)	Barcroft (28) Lazarus-Barlow (173) Blalock (132) Lepore (11) Barbour (94) Underhill (315) Barbour (130)
Tissue effective osmotic pressure	Permeability to proteins Atrophies	Oxygen tension Narcosis Specific ions	Oxygen supply Narcosis Isotonic injections Intradermal fluid	Landis (169) Boycott (49) Gollwitzer-Meyer (119)	 Barbour (21) Hines (142) Aldrich (12)
Blood effective osmotic pressure	Plasma protein concentration Permeability to proteins	Bone-marrow activities Liver activities Kidney activities	Plasmapheresis Low protein diet Vitamine B deficiency Liver injury Renal injury Hypertonic injections Thirst Water drinking	Whipple (291) Frisch (107) Yamaguchi (337) Abe (2) Govaerts (121) Keith (155) Underhill (314) Wilson (336)	Leiter (178,179) Denton (76) Krause (161) Underhill (314) Underhill (314) Tashiro (303)

demonstrated (table 5). These occur in response to changes in the colloid osmotic pressure of the blood, but still more markedly in response to modifications of blood pressure. Moreover, it is known that the changes of water content are not simultaneous in the various tissues, but shifts of concentration occur, first to one tissue and then to another (11). Obviously the water contents of tissue can be understood only in correlation with the entire vasomotor system, a fact that was appreciated by Ludwig (98) and still more by Starling (295, 297).

Many additional measurements of water redistribution have been carried out on sacrificed animals, as is shown in table 6. The skin, particularly, changes its water content with remarkable readiness in

TABLE 5

Functional changes of water content measured in biopsy tissue samples

INFLUENCE	SPECIES	TISSUE	INTERVAL	RESULT	REFERENCE
Isotonic infusion	Dog	Skin	5 minutes	+	Lepore (11)
Starvation	Dog	Skin	3 days	−	Nathan (230)
Isotonic infusion	Dog	Muscle	5 minutes	+	Lepore (11)
Morphine withdrawal	Dog	Muscle	2 days	+	Barbour (23)
Starvation	Dog	Muscle	3 weeks	+	Biddle (39)
Trauma	Dog	Muscle	1 hour	+	Blalock (132)
Anesthetics, adrenalin	Rabbit	Muscle	30 minutes	0	Glass (115)
Decerebration, pituitrin	Rabbit	Muscle	30 minutes	+	Glass (115)
25 per cent NaCl infusion	Rabbit	Muscle	18 minutes	−	Moritz (225)
20 per cent NaCl infusion	Rabbit	Muscle	30 minutes	−	(303, 18)
Caffein	Rabbit	Muscle	30 minutes	−	Sakata (270)
Water (nephrectomized)	Rabbit	Muscle	1 day	+	(303, 18)
Morphine withdrawal	Dog	Liver	2 days	+	Barbour (23)
NaCl infusion	Dog	Liver	5 minutes	+	Lepore (11)

response to a great many stimuli. The only measurements of actual volume changes are those of the skin made along with other tissues by plethysmographs, the changes being invoked usually by venous congestion. The volume of other entire organs, especially the liver, can likewise be measured (202, 97), and shown to change markedly in the presence of numerous chemical and nervous stimuli. How much of such change is due to something other than dilatation of vessels with contained blood cannot, however, usually be decided.

3. *Transudation.* The conditions under which water is poured into or absorbed from the peritoneal cavity have been extensively studied (106, 62, 247). Fluid collects here after intravenous infusions (64), as

well as in response to many forms of peritoneal irritation. The same facts apply in smaller degree to the pleural cavity, and to subdermal spaces.

Measurements of the rate of lymph formation are often convenient indices to various conditions of fluid interchange. The assumption is valid that whenever lymph flow is influenced, the tissue cells with which tissue fluid and lymph are interchanging are likely to be also affected. The lymph formation *per se* is probably a slow process compared to other fluid exchanges within the body; it may be said that lymph formation is important qualitatively but rarely quantitatively.

TABLE 6

Functional changes of water content measured in autopsy tissue samples

INFLUENCE	SPECIES	TISSUES	RESULT	REFERENCE
Starvation	Cat, dog	6	+	Voit (321, 322)
High protein diet	Cat	3	−	(298, 311)
High temperature	Dog	2	+	Hamilton (130)
Pitressin	Rabbit	4	−	Barbour (94)
Cocaine	Rabbit	Liver	+	Barbour (24)
Guanidine	Rabbit	4	+	Ellis (95)
Anesthesia	Rabbit, rat	Brain	+	Barbour (21)
Water deprivation	Dog	10	−	(102, 308)
Water deprivation	Rat	23	−	Kudo (163)
Water deprivation	Rabbit	7	−	Underhill (314)
Burn	Rabbit	5	−	Underhill (315)
Plasma protein depletion	Dog	10	+	Lepore (179)
Infusion	Dog	11	+	(100, 324, 237)
Infusion	Cat	6	+	Skelton (290)
Beriberi, scurvy	Guinea pig	10	0	Krause (161)
Degeneration (cut)	Rabbit	Sciatic Nerve	+	May (204)

When large quantities of salt solution are injected into the circulation the excess fluid leaves the blood relatively rapidly. On dogs and other animals numerous observations have been made to find where the solution is deposited (64, 100, 290, 11). Along with the water, considerable quantities of chloride are deposited in the tissues, and chiefly in the skin (237, 209). While skin takes up most of the fluid per gram of tissue, muscle often holds the larger part of the total mass of fluid. After hemorrhage or dehydration most fluid is drawn from these same tissues. The results on skin are confirmed in man, for after drinking isotonic

sodium chloride solution, visible edema is found in the skin (281, 214, 19).

Pathological accumulations and losses of fluid appear to involve principally these same tissues in cases where generalized factors rather than local factors are responsible. Localized swellings and edemas can be imitated in detail by the application of appropriate suction, vascular, and drug stimuli.

4. *Tissue volumes.* The data on tissues have practically all been gained by measuring the concentration of water relative to its solids or some one constituent. Perhaps equally informing would be the measurement of tissue volumes. This is usually done only relatively by means of plethysmographs and oncometers, and the assumption then must be made that the redistribution of blood in the local vessels is insufficient to account for the changes measured. In this way the effects of blood hypertonicity (329), gravity (hydrostatic pressure) (16, 180, 307, 312, 325, 326), and venous stasis (211, 85) on the volume of particular parts of the body have been measured. These data are only a beginning, and compared to the data that exist on blood volume (263) leave wide gaps in present information.

The various tissues have now been described, as far as the deficient data allow, with respect to their changes of water balances. The same data might be cross-sectioned to show what happens in all parts of the body in response to each sort of stimulus. The only way in which any bodily adjustment can be understood is by picturing the relative adjustments, action and reaction, in each part.

The most important conclusion that emerges is that tissues respond differently in amount and in time, to the same stimulus. Of the chief factors, hydrostatic tensions and effective osmotic tensions, the latter change usually over rather long periods of time, but the former change quickly and are under highly localized control. Obviously, the vasomotor system must now be pictured as quantitatively the largest factor in the distribution of water, overpowering the older view that water was merely a requisite to the maintenance of adequate blood pressure.

D. *Functions of Water in Tissues.* The variety of uses to which water is put in the bodily machinery is as great as the variety of work performed by the tissues themselves.

Muscular contraction undoubtedly involves the translocation of water on a minute scale. Certain elements of the fibers swell while others shrink. Such oriented movement of water may be the limiting factor in the speed of shortening and of lengthening.

Glands in most instances use water as the basis of their products. Here too the active movement of water is under the control of various nervous and pharmacological factors.

Perhaps the most curious functions of water in the body concern the lubrication and protection of organs. In every movable bony joint, synovial fluid is present in regulated amount. Most tendons have similar fluids held in sheaths which surround them. The cerebrospinal fluid acts as a water cushion for the entire central nervous system. Changes in the amount and the pressure of this fluid are of course of the utmost consequence, and the chief factors in producing these changes have been evaluated (33, 327).

In the sense organs aqueous media serve desirable ends. In the eye refraction and optical distances depend on the volumes and pressures of the fluids. In the ear sounds are transmitted from the ear-drum to the sensory end-organs through water. The fluid of the semi-circular canals and of the static organs mediates sensations of motion and gravity. Taste and smell depend upon stimulation by substances that go into solution.

The convection of materials by water is *the* mode of supply to all tissues. Blood, lymph, digestive juices, tears, milk, urine, sweat, and mucus are aqueous. Excretion and absorption, diffusion and secretion all depend upon water. The regulation of the circulation might be considered as a problem in the continual adjustment of the blood volume with respect to a mean position of the vasomotor system and a mean plasma concentration.

It is necessary to refer also to the numerous instances in which tissue function has been shown to change significantly with changes of water content. Isolated muscle and nerve have frequently been studied in this connection, because their responses could be measured quantitatively (87, 88, 57). Studies *in vivo* have been carried out on intestinal motility (258), chronaxie of cutaneous nerves (5), and the responses of various glands. The elasticity of the skin is probably a function of its water content (273, 152). In thirst the secretions of various glands are diminished markedly. The responses of the kidneys to modifications of water content are well known.

E. *Normal Tissue Water Contents.* In addition to the functional changes of water content, there are regular changes of content with growth and other long-time processes. These have been measured in many tissues, particularly brain (330, 63, 278), muscle (126 p. 60, 309), and blood (255) of man, brain of rabbit (17), and muscle of calf (147).

The most extensive data refer to age and sex differences in the tissues of rats (190, 81).

There are diversities between sexes with respect to vapor tension of the blood (195). The differences of water content among species are considerable (227), and experimental or pathological deviations can only be compared closely with control tissues analysed at the same time and in the same manner.

Not only water contents, but protein contents, specific gravities, freezing points, and other data are available for some of the principal tissues such as blood (310, 271, 283, 294), and muscle (148, 173, 126 p. 60).

While these variables have been studied in relation to age and to a few other factors, their correlations and significances are little understood.

The data upon normal tissue composition are today of the most primitive sort; in no tissue except blood are normal values for water content sufficiently established. The interrelations of water content with other measures of hydration are quantitatively almost unknown. The day is still distant when the condition of water balance of any one tissue can be described; even for blood plasma only part of the factors entering the equilibrium have been correlated (140, 275). It is apparent that the problems of water balance can never be settled alone. Only in the light of data concerning the regulation and interrelations of every other constituent and force of the cell and of the body can this familiar substance be fully appreciated.

The author is indebted to Dr. Irvine McQuarrie for contributing valuable references and discussions in the early planning of this review. Some portion of the material was reported to the White House Conference on Child Health and Protection (Growth and Development of the Child, Part III Nutrition, pp. 307–333) at the invitation of Dr. K. D. Blackfan, whose encouragement was a factor leading up to the writing of the present paper. The present paper was completed in December, 1931.

BIBLIOGRAPHY

(1) Abderhalden, E. und B. Bloch. 1907. Zeitschr. physiol. Chem., liii, 464.
(2) Abe, S. 1931. Tohoku Journ. Exp. Med., xvii, 174.
(3) Abe, S. 1931. Tohoku Journ. Exp. Med., xvii, 412.
(4) Abe, K. und S. Sakata. 1925. Arch. Exp. Path. Pharm., cv, 93.
(5) Achelis, J. D. 1930. Pflüger's Arch., ccxxvi, 212.
(6) Adolph, E. F. 1921. Journ. Physiol., lv, 114.
(7) Adolph, E. F. 1923. Amer. Journ. Physiol., lxv, 419.

(8) Adolph, E. F. 1923. Amer. Journ. Physiol., lxvi, 445.
(9) Adolph, E. F. 1930. Quart. Rev. Biol., v, 51.
(10) Adolph, E. F. 1931. Amer. Journ. Physiol., xcvi, 598.
(11) Adolph, E. F. and M. J. Lepore. 1931. Proc. Soc. Exp. Biol. Med., xxviii, 963; also unpublished results.
(12) Aldrich, C. A. and W. B. McClure. 1924. Journ. Amer. Med. Assoc., lxxxii, 1425.
(13) Ambard, L. et E. Papin. 1909. Arch. Internat. Physiol., viii, 437.
(14) Ambard, L. et F. Schmid. 1929. Ann. physiol., v, 393.
(15) Armstrong, H. E. 1908. Proc. Roy. Soc., lxxxi, A, 80.
(16) Atzler, E. und R. Herbst. 1923. Zeitschr. gesammt. exp. Med., xxxviii, 137.
(17) Backlin, E. 1930. Upsala Läkareförenings Förhandlingar, N. F., xxxv, 105.
(18) Baer, R. 1926. Arch. exp. Path. Pharm., cxix, 102.
(19) Baird, M. M. and J. B. S. Haldane. 1922. Journ. Physiol., lvi, 259.
(20) Barbour, H. G. 1921. Physiol. Rev., i, 295.
(21) Barbour, H. G. 1931. Science, lxxiii, 346.
(22) Barbour, H. G. and W. H. Hamilton. 1925. Amer. Journ. Physiol., lxxiii, 665.
(23) Barbour, H. G., L. G. Hunter and C. H. Richey. 1929. Journ. Pharm., xxxvi, 251.
(24) Barbour, H. G. and H. T. Marshall. 1931. Journ. Pharm., xliii, 147.
(25) Barbour, H. G. and E. Tolstoi. 1924. Amer. Journ. Physiol., lxvii, 378.
(26) Barcroft, J. 1900. Journ. Physiol., xxv, 479.
(27) Barcroft, J., H. A. Harris, D. Orahovats and R. Weiss. 1925. Journ. Physiol., lx, 443.
(28) Barcroft, J. and T. Kato. 1916. Phil. Trans., ccvii, B, 149.
(29) Batisweiler, J. 1928. Arch. Gynäkol., cxxxiv, 62.
(30) Bayliss, L. E., and A. R. Fee. 1930. Journ. Physiol., lxx, 60.
(31) Bayliss, W. M. 1920. Journ. Pharm., xv, 29.
(32) Beard, J. W. and A. Blalock. 1931. Arch. Surg., xxii, 617.
(33) Becht, F. C. 1920. Amer. Journ. Physiol., li, 1.
(34) Benedict, F. G. 1915. Carn. Inst. Wash. Publ. no. 203, 416 pp.
(35) Benedict, F. G. and T. M. Carpenter. 1910. Carnegie Inst. Wash. Publ. no. 126, 255 pp.
(36) Benedict, F. G. and R. D. Milner. 1907. U. S. Dept. Agric., O.E.S. Bull. 175, 224.
(37) Benedict, F. G. and H. F. Root. 1926. Arch. Internal Med., xxxviii, 1.
(38) Bernard, C. 1858. Lecons sur la physiologie et la pathologie du système nerveux. Paris. i, 398.
(39) Biddle, H. C. and P. E. Howe. 1913. Biochem. Bull., ii, 386.
(40) Bing, F. C. and L. B. Mendel. 1931. Amer. Journ. Physiol., xcviii, 169.
(41) Bischoff, E. 1863. Zeitschr. rat. Med., Ser. 3, xx, 75.
(42) Bischoff, T. L. W. und C. Voit. 1860. Die Gesetze der Ernährung des Fleischfressers. Leipzig und Heidelberg, 304 S.
(43) Blix, G. 1916. Biochem. Zeitschr., lxxiv, 302.
(44) Blum, L. 1909. Verhandl. Kong. inn. Med., xxvi, 122.

(45) Böhme, A. 1911. Deutsch. Arch. klin. Med., ciii, 522.
(46) Botazzi, F. et al. 1925. H. Wintersteins Handb. vergl. Physiol., 2. Aufl., *1*, i, 1456 S.
(47) Bourquin, H. 1927. Amer. Journ. Physiol., lxxix, 362.
(48) Bourquin, H. 1927. Amer. Journ. Physiol., lxxxiii, 125.
(49) Boycott, A. E., C. Price-Jones. 1922. Journ. Path. Bact., xxv, 335.
(50) Bratusch-Marrain, A. 1928. Monats. Kinderhk., xli, 122.
(51) Bratusch-Marrain, A. 1929. Arch. Kinderhk., lxxxvii, 81.
(52) Bridge, E. M. and E. M. Bridges. 1931. Journ. Biol. Chem., xciii, 181.
(53) Brubacher, H. 1890. Zeitschr. Biol., xxvii, 517.
(54) Brull, L. 1930. Rev. belge sci. méd., ii, 121.
(55) Brunn, F. 1920. Zentralbl. inn. Med., xli, 657.
(56) Bugarszky, S. 1897. Pflüger's Arch., lxviii, 389.
(57) Bürger, M. und L. Lendle. 1925. Arch. exp. Path. Pharm., cix, 1.
(58) Caldwell, G. T. 1931. Physiol. Zool., iv, 324.
(59) Camerer, W., Jr. 1902. Zeitschr. Biol., xliii, 1.
(59a) Carpenter, T. M. and E. L. Fox. 1930. Journ. Nutr., ii, 359.
(60) Carrier, E. B. and P. B. Rehberg. 1923. Skand. Arch. Physiol., xliv, 20.
(61) Chaussin, J. 1920. Journ. physiol. pathol. gen., xviii, 895.
(62) Clark, A. J. 1921. Journ. Pharm., xvi, 415.
(63) Cohn, M. 1907. Deutsch. med. Wochenschr., xxxiii, 1987.
(64) Cohnheim, J. und L. Lichtheim. 1877. Virchow's Arch., lxix, 106.
(65) Cohnheim, O., G. Kreglinger und G. Kreglinger II. 1909. Zeitschr. Physiol. Chem., lxiii, 413.
(66) Cow, D. 1914. Journ. Physiol., xlviii, 1.
(67) Crisler, G. 1928. Amer. Journ. Physiol., lxxxv, 324.
(68) Cushny, A. R. 1926. The secretion of the urine. 2nd ed., London, 288 pp.
(68a) Czerny, A. 1894. Arch. exp. Path. Pharm., xxxiv, 268.
(69) Dale, H. H. and P. P. Laidlaw. 1919. Journ. Physiol., lii, 355.
(70) Daniel, I. und F. Högler. 1922. Wien. Arch. inn. Med., iv, 167.
(71) Darrow, D. C. and T. E. Buckman. 1928. Amer. Journ. Dis. Child., xxxvi, 248.
(72) Dastre, A. et P. Loye. 1889. Arch. physiol. norm. et path., Ser. 5, i, 253.
(73) Dautrebande, L., H. W. Davies and J. Meakins. 1923. Heart, x, 133.
(74) Davies, H. W., J. B. S. Haldane and G. L. Peskett. 1922. Journ. Physiol., lvi, 269.
(75) Dennig, A. 1898. Zeitschr. diatet. physik. Therap., i, 281.
(76) Denton, M. C. and E. A. Kohman. 1918. Journ. Biol. Chem., xxxvi, 249.
(77) De Rudder. 1928. Zeitschr. Kinderhk., xlv, 404.
(78) De Rudder. 1928. Zeitschr. Kinderhk., xlvi, 384.
(79) Dill, D. B., J. H. Talbott and H. T. Edwards. 1930. Journ. Physiol., lxix, 267.
(80) Donaldson, H. H. 1910. Journ. Comp. Neurol., xx, 119.
(81) Donaldson, H. H. 1924. The rat: Data and reference tables. 2nd ed., Philadelphia. 469 pp.
(82) Downs, W. G., Jr. and E. M. K. Geiling. 1929. Proc. Soc. Exp. Biol. Med., xxvii, 63.

(83) Drake, T. G. H., C. F. McKhann and J. L. Gamble. 1930. Journ. Exp. Med., li, 867.
(84) Dresel, K. und Z. Leitner. 1929. Zeitschr. klin. Med., cxi, 394.
(85) Drury, A. N. and N. W. Jones. 1927. Heart, xiv, 55.
(86) Duff, P. A. 1932. Proc. Soc. Exp. Biol. Med., xxix, 508.
(86a) Duke-Elder, W. S. 1927. Journ. Physiol., lxiv, 78.
(87) Durig, A. 1901. Pflüger's Arch., lxxxvii, 42.
(88) Durig, A. 1902. Pflüger's Arch., xcii, 293.
(89) Ebbecke, U. 1917. Pflüger's Arch., clxix, 1.
(90) Ebbecke, U. 1923. Pflüger's Arch., cxcix, 197.
(91) Eckhard, C. 1869. Eckhard's Beitr. Anat. Physiol., iv, 153.
(92) Ederer, S. 1927. Arch. exp. Path. u. Pharm., cxxii, 211.
(93) Elek, L. und E. Róth. 1928. Arch. exp. Path. u. Pharm., cxxxii, 246.
(94) Ellerbrook, G. E., E. S. Dunham and H. G. Barbour. 1930. Journ. Pharm., xxxix, 249.
(95) Ellis, M. M. 1928. Biochem. Journ., xxii, 937.
(96) Elwyn, H. 1929. Edema and its treatment. New York. 182 pp.
(97) Emery, F. E. and F. R. Griffith, Jr. 1931. Journ. Pharm., xlii, 233.
(98) Emminghaus, H. 1873. Arb. physiol. Anst. Leipzig, viii, 51.
(99) Engel, K. und P. Scharl. 1906. Zeitschr. klin. Med., lx, 225.
(100) Engels, W. 1904. Arch. exp. Path. Pharm., li, 346.
(101) Eppinger, H. 1917. Pathologie und Therapie der menschlichen Ödems. Berlin.
(102) Falck, C. P. und T. Scheffer. 1854. Arch. physiol. Heilk., xiii, 508.
(103) Farkas, K. 1908. Landw. Jahrb., xxxvii, 51.
(104) Fehling, H. 1877. Arch. Gynäkol., xi, 523.
(105) Freund, W. 1904. Jahrb. Kinderhk., lix, 421.
(106) Friedemann, U. 1925. Handb. Biochem., 2. Aufl., v, 254.
(107) Frisch, R. A., L. B. Mendel and J. P. Peters. 1929. Journ. Biol. Chem., lxxxiv, 167.
(108) Fujimaki, Y. und F. Hildebrandt. 1924. Arch. exp. Path. u. Pharm., cii, 226.
(109) Galeotti, G. und E. Signorelli. 1912. Biochem. Zeitschr., xli, 268.
(110) Gamble, J. L., M. C. Putnam and C. F. McKhann. 1929. Amer. Journ. Physiol., lxxxviii, 571.
(111) Gänsslen, M. 1922. Münch. med. Wochenschr., 263.
(112) Gasser, H. S., J. Erlanger and W. J. Meek. 1919. Amer. Journ. Physiol., l, 31.
(113) Gerhartz, H. 1910. Pflüger's Arch., cxxxv, 104.
(114) Gildemeister, M. und L. Hoffmann. 1922. Pflüger's Arch., cxcv, 153.
(115) Glass, A. 1928. Arch. exp. Path. u. Pharm., cxxxvi, 72.
(116) Goldschmidt, S. 1921. Physiol. Rev., i, 421.
(117) Gollwitzer-Meier, K. 1928. Pflüger's Arch., ccxviii, 586.
(118) Gollwitzer-Meier, K. und W. Bröcker. 1928. Zeitschr. gesammt. exp. Med., lxii, 105.
(119) Gollwitzer-Meier, K. und R. Rabl. 1926. Zeitschr. gesammt. exp. Med., liii, 525.
(120) Gottschalk, N. 1926. Pflüger's Arch., ccxii, 441.

(121) Govaerts, P. 1924. Bull. Acad. Roy. Méd. Belg., Ser. 5, iv, 161.
(122) Govaerts, P. 1927. Bull. Acad. Roy. Méd. Belg., Ser. 5, vii, 356.
(123) Grawitz, E. 1895. Berl. klin. Wochenschr., xxxii, 713, 740.
(124) Greene, C. H. and L. G. Rowntree. 1927. Amer. Journ. Physiol., lxxx, 209.
(124a) Grollman, A. 1929. Amer. Journ. Physiol., lxxxix, 157.
(125) Grollman, A. 1931. Journ. Gen. Physiol., xiv, 661.
(126) Gundobin, N. P. 1911. Die Besonderheiten des Kindersalters. Berlin. 592 S.
(127) Haake, B. und K. Spiro. 1902. Beitr. chem. Physiol., ii, 149.
(128) Haldi, J., J. Larkin and P. Wright. 1929. Amer. Journ. Physiol., lxxxviii, 112.
(129) Hamburger, H. J. 1902. Osmotischer Druck und Ionenlehre. Bände 3. 1902–1904.
(130) Hamilton, W. F. and H. G. Barbour. 1925. Amer. Journ. Physiol., lxxiii, 321.
(131) Hammett, F. S. 1927. Endocrinol., xi, 297.
(132) Harris, P. N. and A. Blalock. 1931. Arch. Surg., xxii, 638.
(133) Hatai, S. 1917. Amer. Journ. Anat., xxi, 23.
(134) Hecht, A. F. 1924. Zeitschr. Kinderhk., xxxviii, 192.
(135) Heilig, R. 1924. Klin. Wochenschr., iii, 1117.
(136) Heilner, E. 1906. Zeitschr. Biol., xlvii, 538.
(137) Heller, H. und A. Schwarz. 1930. Zeitschr. ges. exp. Med., lxx, 760.
(138) Heller, H. und A. Schwarz. 1930. Zeitschr. ges. exp. Med., lxxi, 416.
(139) Henderson, L. J. 1913. The fitness of the environment. New York. 317 pp.
(140) Henderson, L. J. 1928. Blood, a study in general physiology. New Haven. 397 pp.
(141) Hill, A. V. and P. S. Kupalov. 1930. Proc. Roy. Soc., cvi, B, 445.
(142) Hines, H. M., C. E. Leese and H. R. Jacobs. 1927. Amer. Journ. Physiol., lxxxiii, 269.
(143) Houghten, F. C., W. W. Teague, W. E. Miller and W. P. Yant. 1929. Amer. Journ. Physiol., lxxxviii, 386.
(144) Istamanowa, T. und Chudoroschewa. 1930. Zeitschr. gesammt. exp. Med., lxxi, 212.
(145) Ivy, A. C. 1918. Amer. Journ. Physiol., xlvi, 340.
(146) Jackson, C. M. and V. D. E. Smith. 1931. Amer. Journ. Physiol., xcvii, 146.
(147) Jacubowitsch, W. 1892. Arch. Kinderhlk., xiv, 355.
(148) Jappelli, A. 1906. Arch. int. physiol., iv, 369.
(149) Jarisch, A. und G. Liljestrand. 1927. Skand. Arch. Physiol., li, 235.
(150) Jores, A. 1930. Zeitschr. gesammt. exp. Med., lxxi, 170.
(151) Jürgensen, E. 1924. Deutsch. Arch. klin. Med., cxliv, 193.
(152) Katz, J. R. 1917. Koll.-chem. Beih., ix, 1.
(153) Keeton, R. W. 1924. Amer. Journ. Physiol., lxix, 307.
(154) Keith, N. M. 1922. Amer. Journ. Physiol., lix, 452.
(155) Keith, N. M. 1924. Amer. Journ. Physiol., lxviii, 80.
(156) Klein, O. und W. Nonnenbruch. 1930. Zeitschr. klin. Med., cxii, 568.

(157) KLEITMAN, N. 1927. Amer. Journ. Physiol., lxxxi, 336.
(158) KNAFFL-LENZ, E. UND S. NOGAKI. 1925. Arch. exp. Path. Pharm., cv, 109.
(159) KOHMAN, E. A. 1920. Amer. Journ. Physiol., li, 378.
(160) KORANYI, A. 1897. Zeitschr. klin. Med., xxxiii, 1; xxxiv, 1.
(161) KRAUSE, D. J. 1922. Amer. Journ. Physiol., lx, 234.
(161a) KROGH, A. 1922. The anatomy and physiology of capillaries. New Haven. 276 pp. Also 2nd edition, 1929.
(162) KUDO, T. 1921. Amer. Journ. Anat., xxviii, 399.
(163) KUDO, T. 1921. Journ. Exp. Zool., xxxiii, 435.
(164) KUNITZ, M. 1926. Journ. Gen. Physiol., ix, 715.
(165) KUNITZ, M. 1927. Journ. Gen. Physiol., x, 811.
(166) LAMSON, P. D. AND J. ROCA. 1921. Journ. Pharm., xvii, 481.
(167) LAMY, H. ET A. MAYER. 1906. Journ. Physiol., path. gén., viii, 258.
(168) LANDIS, E. M. 1927. Amer. Journ. Physiol., lxxxii, 217.
(169) LANDIS, E. M. 1928. Amer. Journ. Physiol., lxxxiii, 528.
(170) LANDIS, E. M. 1930. Amer. Journ. Physiol., xciii, 353.
(171) LANDIS, E. M. 1930. Heart, xv, 209.
(172) LANDIS, E. M. 1931. Amer. Journ. Physiol., xcviii, 704.
(173) LAZARUS-BARLOW, W. S. 1894. Phil. Trans.. clxxxv, B, 779.
(174) LAZARUS-BARLOW, W. S. 1895. Journ. Physiol., xix, 140.
(175) LAZARUS-BARLOW, W. S. 1896. Journ. Physiol., xix, 418.
(176) LAZARUS-BARLOW, W. S. 1896. Journ. Physiol., xx, 145.
(177) LEE, F. W., E. B. CARRIER AND G. H. WHIPPLE. 1922. Amer. Journ. Physiol., lxi, 149.
(178) LEITER, L. 1931. Arch. Internal Med., xlviii, 1.
(179) LEPORE, M. J. 1932. Arch. Internal Med., l, 488.
(180) LEWIS, T. 1924. Heart, xi, 119.
(181) LEWIS, T. 1927. The blood vessels of the human skin and their responses. London, 322 pp.
(182) LICHTWITZ, L. 1911. Arch. exp. Path. Pharm., lxv, 128.
(185) LINDEMANN, L. 1900. Deutschr. Arch. klin. Med., lxv, 1.
(184) LOEWEN, D. F., M. E. FIELD AND C. K. DRINKER. 1931. Amer. Journ. Physiol., xcviii, 70.
(185) LOEWY, A. 1914. Biochem. Zeitschr., lxvii, 243.
(186) LOEWY, A. 1926. Tab. Biol., iii, 506.
(187) LOEWY, A. UND H. GERHARTZ. 1913. Pflüger's Arch., clv, 231.
(188) LOEWY, A. UND W. WECHSELMANN. 1911. Virchow's Arch., ccvi, 79.
(189) LOMBARD, W. P. 1906. Journ. Amer. Med. Assoc., xlvii, 1790.
(190) LOWREY, L. G. 1913. Anat. Rec., vii, 143.
(190a) LUBLIN, A. 1929. Zeitschr. klin. Med., cix, 371.
(191) LUSK, G. 1912. Journ. Biol. Chem., xiii, 27.
(192) MAGNUS, R. 1900. Arch. exp. Path. Pharm., xliv, 396.
(193) MAGNUS, R. 1901. Arch. exp. Path. Pharm., xlv, 210.
(194) MANCHESTER, R. C., C. HUSTED AND I. MCQUARRIE. 1931. Journ. Nutr., iv, 39.
(195) MARGARIA, R. 1930. Journ. Physiol., lxx, 417.
(196) MARRIOTT, W. M. 1923. Physiol. Rev., iii, 275.
(197) MARSHALL, E. K., JR. 1920. Journ. Pharm., xvi, 141.

(198) Marshall, H. T., B. F. Aydelotte and H. G. Barbour. 1931. Amer. Journ. Physiol., xcviii, 615.
(199) Marx, H. 1926. Deutsch. Arch. klin. Med., clii, 354.
(200) Marx, H. 1926. Deutsch. Arch. klin. Med., cliii, 358.
(201) Marx, H. und W. Mohr. 1927. Arch. exp. Path. Pharm., cxxiii, 205.
(202) Mattson, H. 1929. Amer. Journ. Physiol., xc, 146.
(203) Mautner, H. 1927. Arch. exp. Path. Pharm., cxxvi, 255.
(204) May, R. M. 1930. Bull. soc. chim. biol., xii, 934.
(205) Mayer, A. et G. Nichita. 1930. Ann. physiol., v, 774.
(206) McClendon, J. F. 1931. Amer. Journ. Physiol., xcviii, 216.
(207) McIntosh, R., L. Kajdi and D. Meeker. 1930. Journ. Clin. Inves., ix, 333.
(208) McQuarrie, I. 1929. Amer. Journ. Dis. Child., xxxviii, 451.
(209) Melli, G. et G. Tasso. 1930. Rev. belge sci. méd., ii, 209.
(210) Mello-Leitao. 1916. Amer. Journ. Dis. Child., xi, 214.
(211) Mende. 1919. Deutsch. Zeitschr. Chir., cl, 379.
(212) Meyer, J. 1925. Rev. franç. derm. vénéréol., i, 249; 320.
(213) Meyer, J. 1925. Rev. franç. pédiat., i, 409.
(214) Meyer, L. F. und S. Cohn. 1911. Zeitschr. Kinderhk., ii, 360.
(215) Meyer-Bisch, R. und W. Wohlenberg. 1926. Zeitschr. gesammt. exp. Med., l, 728.
(216) Michel, C. 1900. L'Obstétrique, v, 252.
(217) Molitor, H. 1928. Wien. klin. Wochenschr., xli, 194.
(218) Molitor, H. und E. Pick. 1925. Arch. exp. Path. Pharm., cvii, 180; 185.
(219) Moog, O. und K. Buchheister. 1926. Münch. med. Wochenschr., lxxiii, 895.
(220) Moog, O. und E. T. Nauck. 1921. Zeitschr. gesammt. exp. Med., xxv, 385.
(221) Moore, N. S. and H. J. Stewart. 1930. Journ. Clin. Inves., ix, 423.
(222) Moraczewski, W. 1924. Zeitschr. klin. Med., ci, 38.
(223) Moraczewski, W. 1925. Journ. Physiol. Path. Gen., xxiii, 12.
(224) Moran, T. 1930. Proc. Roy. Soc., cvii, B, 182.
(225) Moritz, F. 1887. Deutsch. Arch., klin. Med., xli, 395.
(226) Moss, K. N. 1923. Proc. Roy. Soc., xcv, B, 181.
(227) Moulton, C. R. 1923. Journ. Biol. Chem., lvii, 79.
(228) Mudd, S. 1926. Journ. Gen. Physiol., ix, 361.
(229) Nakashima, K. 1924. Journ. Biochem., iv, 277.
(230) Nathan, E. und F. Stern. 1931. Dermatol. Zeitschr., lx, 299.
(231) Nelson, O. A. and G. A. Hulett. 1920. Journ. Ind. Engin. Chem., xii, 40.
(232) Newburgh, L. H., M. W. Johnston and M. Falcon-Lesses. 1930. Journ. Clin. Invest., viii, 161.
(233) Newton, R. and R. A. Gortner. 1922. Bot. Gaz., lxxiv, 442.
(234) Niemann, A. 1915. Jahrb. Kinderhk., lxxxii, 21.
(235) Nonnenbruch, W. 1926. Handb. norm. path. Physiol., xvii, iii, 223.
(236) Onohara, K. 1925. Biochem. Zeitschr., clx, 427.
(237) Padtberg, J. H. 1910. Arch. exp. Path. Pharm., lxiii, 60.
(238) Parnas, J. K. 1926. Handb. norm. path. Physiol., xvii, iii, 137.

(239) Pavlov, I. P. 1910. The work of the digestive glands. 2nd ed., London 266 pp.
(240) Petrilli, A. Unpublished.
(241) Pfeiffer, E. 1911. Verhandl. Kong. inn. Med., xxviii, 506.
(242) Pick, E. P. 1928. Deutsch. Zeitschr. Nervenhk., cvi, 238.
(243) Pilcher, J. D. 1926. Amer. Journ. Dis. Child., xxxi, 77.
(244) Poulsson, L. T. 1930. Zeitschr. gesammt. exp. Med., lxxi, 577.
(245) Priestley, J. G. 1921. Journ. Physiol., lv, 305.
(246) Pugliese, A. 1910. Zeitschr. Biol., liv, 100.
(247) Putnam, T. J. 1923. Amer. Journ. Physiol., lxiii, 548.
(248) Rabinovitsch, J. 1927. Amer. Journ. Physiol., lxxxii, 279.
(249) Regnier, A. 1916. Zeitschr. exp. Path. Therap., xviii, 139.
(250) Ribadeau-Dumas, L. et D. Tisserand. 1926. Nourisson, xiv, 79.
(251) Richardson, H. B. 1926. Journ. Biol. Chem., lxvii, 397.
(252) Richter, C. P. 1930. Brain, liii, 76.
(253) Richter, C. P. and M. E. Brailey. 1929. Amer. Journ. Physiol., xc, 494.
(254) Robinson, W. 1931. Journ. Biol. Chem., xcii, 699.
(255) Rominger, E. 1920. Zeitschr. Kinderhk., xxvi, 23.
(256) Rominger, E. 1928. Monatschr. Kinderhk., xli, 56.
(257) Rose, W. B., C. J. Stucky and L. B. Mendel. 1931. Amer. Journ. Physiol., xci, 520.
(258) Rose, W. B., C. J. Stucky and G. R. Cowgill. 1930. Amer. Journ. Physiol., xci, 531.
(259) Rosemann, R. 1927. Handb. norm. path. Physiol., iii, 819.
(260) Rowntree, L. G. 1922. Physiol. Rev., ii, 116.
(261) Rowntree, L. G. 1923. Arch. Internal Med., xxxii, 157.
(262) Rowntree, L. G. 1926. Journ. Pharm., xxix, 135.
(263) Rowntree, L. G. and G. E. Brown. 1929. The volume of the blood and plasma. Mayo Clinic Monog., Philadelphia. 219 pp.
(264) Rubner, M. 1893. Arch. Hyg., xvi, 101.
(265) Rubner, M. 1898. Arch. Hyg., xxxiii, 151.
(266) Rubner, M. 1902. Die Gesetze des Energieverbrauchs bei der Ernährung. Leipzig und Wien. 426 S.
(267) Rubner, M. 1922. Abhandl. preus. Akad. Wiss., 1.
(268) Rupp, H. 1928. Zbl. Gynäkol., lii, 1119.
(269) Rusznyak, S. 1924. Zeitschr. gesammt. exp. Med., xli, 532.
(270) Sakata, S. 1925. Arch. exp. Path. Pharm., cv, 11.
(271) Salge, B. 1911. Zeitschr. Kinderhk., ii, 347.
(272) Schaal, H. 1922. Biochem. Zeitschr., cxxxii, 295.
(273) Schade, H. 1912. Zeitschr. exp. Path. Therap., xi, 369.
(274) Schade, H. 1924. Handb. Biochem., 2 Aufl, viii, 149.
(275) Schade, H. 1927. Ergebn. inn. Med., xxxii, 425.
(276) Schade, H. und F. Claussen. 1924. Zeitschr. klin. Med., c, 363.
(277) Schade, H. und H. Menschel. 1923. Zeitschr. klin. Med., xcvi, 279.
(278) Schiff, E. und E. Stransky. 1921. Jahrb. Kinderhk., xcvi, 245.
(279) Schittenhelm, A. und H. Schlecht. 1919. Zeitschr. gesammt. exp. Med., ix, 1.
(280) Schlieper, C. 1930. Biol. Rev., v, 309.

(281) Schloss, E. 1910. Jahrb. Kinderhk., lxxi, 296.
(282) Schlüter, E. 1925. Pflüger's Arch., ccx, 432.
(283) Schulz, P. 1912. Zeitschr. Kinderhielk., iii, 495.
(284) Serebrijski, I. und H. Vollmer. 1925. Biochem. Zeitschr., clxiv, 1.
(285) Schackell, L. F. 1909. Amer. Journ. Physiol., xxiv, 325.
(286) Shelburne, S. A. and W. C. Egloff. 1931. Arch. Int. Med., xlvi..
(287) Sherrington, C. S. and S. M. Copeman. 1893. Journ. Physiol., xiv,
(288) Siebeck, R. 1926. Handb. norm. path. Physiol., xvii, iii, 161.
(289) Siebeck, R., und J. Borkowski. 1919. Deutsch. Arch. klin. Med., cx
55.
(290) Skelton, H. 1927. Arch. Int. Med., xl, 140.
(291) Smith, H. P., A. E. Belt and G. H. Whipple. 1920. Amer. Journ. Physi..,
lii, 54.
(292) Snell, A. M. and L. G. Rowntree. 1928. Amer. Journ. Physiol., lxxxv,
577.
(293) Soderstrom, G. F. and E. F. DuBois. 1917. Arch. Int. Med., xix, 931.
(294) Spencer, H. 1929. Amer. Journ. Dis., Child., xxxvii, 546.
(295) Starling, E. H. 1896. Journ. Physiol., xix, 312.
(296) Starling, E. H. 1899. Journ. Physiol., xxiv, 317.
(297) Starling, E. H. 1909. The fluids of the body. London. 186 pp.
(298) Straub, W. 1899. Zeitschr. Biol., xxxviii, 537.
(299) Strauss, H. 1903. Ther. Gegenw., 1903, 192.
(300) Strauss, W. und C. Müller. 1929. Zeitschr. Hyg., cx, 413.
(301) Sutherland, G. F. 1921. Amer. Journ. Physiol., lv, 258.
(302) Svedberg, T. 1924. Journ. Amer. Chem. Soc. xlvi, 2673.
(303) Tashiro, N. 1926. Arch. exp. Path. u. Pharm., cxi, 218.
(304) Thoenes, F. 1925. Biochem. Zeitschr., clvii, 174.
(305) Thomas, K. 1911. Arch. Physiol., 9.
(306) Thompson, W. H. 1900. Journ. Physiol., xxv, 487.
(307) Thompson, W. O., P. K. Thompson and M. E. Dailey. 1928. Journ.
Clin. Inves., v, 573.
(308) Tobler, L. 1910. Arch. exp. Path. u. Pharm., lxii, 431.
(309) Tobler, L. 1911. Jahrb. Kinderhk., lxxiii, 566.
(310) Toyama, I. 1919. Journ. Biol. Chem., xxxvii, 161.
(311) Tsuboi, J. 1903. Zeitschr. Biol., xliv, 376.
(312) Turner, A. H., M. I. Newton and F. W. Haynes. 1930. Amer. Journ.
Physiol., xciv, 507.
(313) Underhill, F. P., G. L. Carrington, R. Kapsinow and G. T. Pack. 1923.
Arch. Int. Med., xxxii, 31.
(314) Underhill, F. P. and M. E. Fisk. 1930. Amer. Journ. Physiol., xcv, 348.
(315) Underhill, F. P., M. E. Fisk and R. Kapsinow, 1930. Amer. Journ.
Physiol., xcv, 339.
(316) Underhill, F. P. and M. A. Sallick. 1925. Journ. Biol. Chem., lxiii,
61.
(317) Veil, W. H. 1916. Deutsch. Arch. klin. Med., cxix, 376.
(318) Veil, W. H. 1918. Biochem. Zeitschr., xci, 267.
(319) Veil, W. H. 1923. Ergebn. inn. Med., xxiii, 648.
(320) Veit, J. 1900. Zeitschr. Geburts., xlii, 316.

(321) VOIT, C. 1866. Zeitschr. Biol., ii, 307.
(322) VOIT, C. 1894. Zeitschr. Biol., xxx, 510.
(323) VOLHARD, F. 1931. Handb. inn. Med., 2. Aufl., vi, i. 1023 S.
(324) WAHLGREN, V. 1909. Arch. exp. Path. u. Pharm., lxi, 47.
(325) WATERFIELD, R. L. 1931. Journ. Physiol., lxxii, 110.
(326) WATERFIELD, R. L. 1931. Journ. Physiol., lxxii, 121.
(327) WEED, L. H. 1922. Physiol. Rev., ii, 171.
(328) WEED, L. H. AND P. S. MCKIBBEN. 1919. Amer. Journ. Physiol., xlviii, 512.
(329) WEED, L. H. AND P. S. MCKIBBEN. 1919. Amer. Journ. Physiol., xlviii, 531.
(330) WEISBACH, A. 1868. Med. Jahrb., xvi, 46.
(330a) WELLS, H. S. 1931. Amer. Journ. Physiol., xcix, 209.
(331) WEYRICH, V. 1862. Die unmerkliche Wasserverdunstung der menschlichen Haut. Leipzig. 250 S.
(332) WEYRICH, W. 1865. Beobachtungen über die unmerkliche Wasserausscheidung der Lungen und ihr Verhältniss zur Hautperspiration. Inaug.-Diss. Dorpat. 46 S.
(333) WIDAL, F. ET A. JAVAL. 1903. Bull. Mém. soc. méd. Hôp. Paris, Ser. 3, xx, 733.
(334) WILEY, F. H. AND L. H. NEWBURGH. 1931. Journ. Clin. Inves., x, 723.
(335) WILLS, F. AND P. B. HAWK. 1914. Journ. Amer. Chem. Soc., xxxvi, 158.
(336) WILSON, T. M. 1905. Amer. Journ. Physiol., xiii, 139.
(337) YAMAGUCHI, T. 1928. Tohoku Journ. Exp. Med., x, 435.
(338) ZANGEMEISTER, W. 1903. Zeitschr. Geburts., l, 385.
(339) ZANGEMEISTER, W. 1919. Zeitschr. Geburts., lxxxi, 491.
(340) ZUNTZ, N., A. LOEWY, F. MÜLLER AND W. CASPARI. 1906. Höhenklima und Bergwanderung in ihrer Wirkung auf den Menschen. Berlin usw. 494 S.

THE ACID OUTPUT OF THE KIDNEY AND THE SO-CALLED ALKALINE TIDE

CHARLES E. BRUNTON

Department of Physiology, London Hospital Medical College, University of London

SECTION I. URINARY ACIDITY AND ACID OUTPUT. The conception of an alkaline tide is usually attributed to Bence Jones (1845, 1849) who observed an "alkalescence of the urine" after meals when he tested the urine by titration to the turning point of litmus. Roberts (1859), whose article contains almost all the knowledge we possess today and much that has been forgotten, used the term "alkaline tide" for this decreased titratable acidity of the urine after meals. The phrase has also been used to describe increases of pH (Munford and Hubbard, 1926; Kaye, 1929); to describe increases in the ratio of urinary B_2HPO_4/BH_2PO_4 (Leathes, 1919; Campbell, 1920; Watson, 1925; Small, 1928) and a decreased output in unit time of acid or of acid plus ammonia (Dodds, 1923).

This article is written from the point of view that, however interesting may be temporary changes of titratable acidity or of pH, they cannot in the nature of things, be correlated mathematically with such quantities as the volume of expired CO_2 or of gastric juice and so are of little use in metabolic studies. The time has come to ask for full information about the individual urinary constituents (Cl, SO_4, PO_4, Na, K, Ca, Mg, NH_4, the organic acids and other organic substances). Such a study has been attempted by Goiffon (1931) and Chatron (1932) on 24 hour specimens and forms one method of calculating the "base economy," a conception studied by Gamble, Ross and Tisdall (1923).

When the reaction of the urine is brought to pH 8.2 (the turning point of phenolphthalein) by titration with NaOH the ions so neutralised have been regarded as coming from the acid phosphate (Henderson, 1911; Henderson and Palmer, 1914). Lactic acid may be present in the urine during severe exercise and in conditions of alkalemia (Macleod and Knapp, 1919). Carbonates also appear in alkaline urines. Betahydroxybutyric acid and acteoacetic acid are found in cases of diabetes and starvation. A method has been devised by Van Slyke and Palmer (1920) for the estimation of total organic acids of urine based on their buffering

power after carbonates and phosphates have been removed. Subsequent modifications in it have been discussed by Peters and Van Slyke (1932) and by Chatron (1932). If, owing to the small volume of urine passed in a 15 or 30 minute period, a determination of the separate constituents cannot as yet be attempted, one should at least know the volumes of urine produced in unit time. When figures are published which merely show a decrease in the titratable acidity of the urine during the 2nd, 3rd and 4th half-hour periods after a vaguely described meal, the decrease may at first suggest a correlation with output of gastric juice. Yet if the volumes of the urine are measured as well, their increase during the periods mentioned (owing to the fluid of the meal), may, when multiplied by the titratable acidities, prove that the acid output actually rose instead of falling. As D. W. Wilson (1923) wrote: "The reaction of the urine, while an index of the output of acid, does not furnish much information about the quantity excreted." In evaluating the usefulness of different articles to be described in this review the reader should discriminate between those which only report changes in the titratable acidity or the pH and those which include information about the minute acid output.

It is the function of the kidneys to help in the regulation of the osmotic pressure and the reaction of the blood as well as to excrete (in solutions where their concentrations are limited) substances which are present to excess in the blood while retaining those which are useful. For the removal of excess acid the kidney has the help of the respiratory system which can increase the output of CO_2 by the lungs. During alkalemia the centre may have the power also to reduce the CO_2 loss. It is to be expected that a number of different causes will produce increased alkalinity of the urine rather than that a single cause should operate in all cases. In consequence, this review has been arranged under headings which deal with various influences on the urinary acid.[1] After a reference to the proposed criteria of an alkaline tide and an account of methods for estimating urinary acid the sections deal with various possible factors which may influence the acidity or acid output of urine:—the digestive secretions and especially the gastric juice, the fluid intake, food constituents (which may act through gastric stimulation, through their inorganic radicles, through preferential absorption in time of certain radicles by the intestine, through the metabolism of organic acids, through the influence of certain foods as "protein savers"

[1] The term "urinary acid" is used throughout this article in a loose sense which covers both "titratable acidity" and "acid output in unit time."

or through specific effects on the nervous system or kidneys), changes in the respiratory centre and wakening from sleep.

The acidity of specimens and the acid output have never been related to a standard acidity or output. If the average titratable acidity of 24 hour specimens, carefully collected from normal adults on a mixed diet, was multiplied by the average 24-hour volume an average daily total acid output would be obtained from which an average hourly (or minute) volume of 0.1 N. acid could be calculated. During the 24 hours the minute acid output would vary around this average level. An alkaline tide could be defined in relation to such a level but would be apt to give us an unjustifiable feeling of knowledge in spite of our real ignorance of some factors in the individual observations. The term "alkaline tide" should be retained as it was originally used (Roberts, 1859): for the phenomenon of decreased titratable acidity of urine after meals, or possibly extended to include increase of urinary pH (Munford and Hubbard, Kaye and others). Used in this sense, however, the alkaline tide becomes an historically interesting but "pot-bound" fruitless observation which cannot be related to other changes in the body as the minute acid-ammonia output might be. Physiologists and clinicians alike should realise its limitations. It is suggested that they should adopt the conceptions of "decreased (minute) acid output" or "decreased minute acid-ammonia output" which would include not only the acid or acid and ammonia content of the samples but also their volume in unit time.

Section II. Methods for estimating urinary acid. The importance of accurate measurement of small volumes of urine collected during such short periods as 15 or 30 minutes is obvious but has seldom been stated. Serious errors in the calculated output may result from imperfect emptying of the bladder (which is easy if the subject is preoccupied) or owing to unstandardised glassware. The minute output of urine may well be less than 0.5 cc. per minute.

Bence Jones (1845) first used the reaction of the urine to litmus and later its titration value to litmus with sodium carbonate as a test of its acidity. Titration with 0.1 N. NaOH to the end point of phenolphthalein has been the usual method in England, modified by the addition of potassium oxalate as suggested by Folin (1903). Henderson (1911) discussed very fully the estimation of urinary acidity and gave methods for examining the urine both in this article and in a later one with Palmer (Henderson and Palmer, 1913a). In the first article he points out the incompleteness of the chemical reactions at the end

points of methyl orange and phenolphthalein as used to measure titratable acidity (p. 409). Results obtained with phenolphthalein, he thinks, serve a useful purpose when compared among themselves but "cannot be utilised for the study of the last step of intermediary metabolism, the regulating action of the kidney, if for no other reason because the blood has considerable 'acidity' as measured by this indicator. Certainly," he continues, "such results measure neither the intensity of the acidity of the urine nor the excess of acid there present as compared with the blood" (p. 410). He declares that the minute volume of urine excreted multiplied by the sum of the figures obtained by titrating the urine to the pH of blood and that representing the ammonia content of the urine in terms of equivalent acid, i.e., Min. Vol $\times$ (Acid + NH_4) content "measures the total excess of acid above the normal amount present in the blood and there associated with the urinary base which has been removed by the kidney. It measures also the total retention of alkali because an equivalent of alkali has been separated from the acid by the kidney and restored to the circulating blood" (p. 412) . . . . "The excess of acid above that present in blood may be associated either with a large or a small amount of saline material consisting of acids and bases in the proportion found in blood. Thus the efficiency of the process varies with the relative and not with the absolute excess of acid over basic material in the urine; it is in fact only to be estimated by . . . the measurement of the actual acidity or hydrogen ion concentration." The human physiologist will certainly be dissatisfied with a standard of efficiency for the kidney based merely on its ability to excrete the maximum number of hydrogen ions with a minimum amount of salt but this pernicious conception of the value of the urinary pH by itself appears to have hidden the more useful conception of the minute acid excretion which has been quoted above from the same author. This eclipse is exemplified in the fate of the methods of urinary examination described in this article and in a subsequent one (Henderson and Palmer, 1913a). In the first method (Henderson, 1911, and Henderson and Palmer, 1914) 25 cc. of urine were diluted with water in a 250 cc. graduated flask, using a small quantity of potassium oxalate. The precipitate of calcium oxalate was removed, equal quantities of neutral red were added to the diluted urine and to a phosphate buffer having the pH of blood and the urine was titrated with 0.1 N. NaOH till the colours matched. This titration value to the pH of blood, added to the acid equivalent of an ammonia estimation could be multiplied by the minute volume of the urine so as to measure the excess of acid which the kidney

removes. In the second method (Henderson and Palmer 1913a) buffer solutions of different pHs were made up and tinted to match the colour of the urine samples. Equal amounts of indicator solutions having been added to the tinted buffer solutions and to the urine diluted with water, the pH of the urine was ascertained by colour matching to within 0.1 pH unit. The urine was not titrated to the pH of blood and so no figure was obtained which might be multiplied by the minute volume. The method however enables so many specimens to be examined in a short time that it was adopted by almost all workers in the United States (Myers and Muntwyler, 1928). Henderson's method of 1911 was forgotten. Palmer, Salvesen and Jackson (1920–21) decreased the dilution of the Henderson and Palmer method to one part in five and introduced phenol red for the range 4.7 to 6.3 pH units. Fiske (1921a) employed four indicators of the sulphonephthalein type for different ranges. He found that a fall of pH might occur after meals without much fall in the titratable acidity. He did not use the volumes of the specimens to relate the urinary acid to gastric secretion. Myers and Booher (1924–25) employed two wedges of dilute indicators of different pHs (and sometimes a third wedge to compensate for turbidity) in a bicolorimeter (Myers, 1922) which allowed the urine to be covered with oil (see Marshall below). Myers and Muntwyler (1928) studied the effects of temperature, dilution, salt content and carbon dioxide on the method and altered the technique so as to dilute the indicators with 0.9 per cent NaCl, to standardise the eight indicator wedges and to correct the result from laboratory temperature of about 20°C. to 38°C. They claim an accuracy of 0.1 pH units at 38°C. Hastings, Sendroy and Robson (1925) had claimed the same accuracy for their method which was carried out in a water bath at 38°C., the urine being under oil.

Of those who have used a change of pH as indicative of an alkaline tide, Munford and Hubbard (1926) stated that "if any specimen is more alkaline than one collected earlier in the test by 1.0 pH or more, or if any two specimens are more alkaline than an earlier one by between 0.5 and 1.0 pH a tide is present. If the average reaction of all sepecimens is alkaline, i.e., is greater than 7.0 pH, the test should be disregarded." Kaye (1929) accepted as evidence of an alkaline tide "in samples of urine taken hourly a rise of 1.0 pH unit in reaction as compared with that of the preceding sample or an increase of 0.5–1.0 pH unit in each of two successive samples as compared with the pH of the urine collected one hour before the first of these samples." Kaye used the quinhydrone electrode in most of his work. None of those who measured

the urinary pH seem to have realized how sterile a conception it is for metabolic studies and how fallacious may be the quick conclusions that one is tempted to draw from it as regards other functions of the body than the kidney. Two unjustifiable assumptions are often made:—first, that the pH is always proportional to the titratable acidity and that, therefore, an increase of pH means a decreased excretion of acid ions; and, secondly, that the volume of urine excreted in unit time is constant so that variation in the urinary acidity proves the existence of correlated variations in the loss of acid ions from the body. As Fiske (1921a) showed, no constant relationship can be predicted between the pH and the titratable acidity of the urine (though their curves are often parallel) since it will vary with the presence of different buffer substances: phosphates, carbonates, organic acids, etc. Further, titration to pH 8.2 will give different results from titration to pH 7.4. Henderson (1911) pointed out that if 0.1 M. NaH_2PO_4 and a mixture of 0.01 M. NaH_2PO_4 and 0.05 M. Na_2HPO_4 are titrated to phenolphthalein both give the same titration figure, while, if titrated to the pH of blood the former is 0.009 N. and the latter 0.003 N.

Titration with 0.1 N. NaOH to phenolphthalein was used by English workers before Leathes' article (1919). Leathes titrated the urine with 0.1 N. alkali to phenolphthalein and with 0.1 N. acid to methyl orange. He stated that the ratio of these two titration numbers was usually that of NaH_2PO_4 to Na_2HPO_4 and was proportional to the reaction of the urine. He called the titration figure of $Na_2HPO_4 \times 100/$ sum of the titrations the "percentage alkalinity." In his results he reported the volumes of urine passed in unit time but merely to show the absence of a diuresis. The product of a minute volume and a pH would have no useful significance. Calvert, Mayrs and Milroy (1921) confirmed the proportionality between the "percentage alkalinity" and the pH in urines from a number of normal and diabetic subjects but unfortunately the parallelism does not always hold. The sum of Leathes' two titrations includes bicarbonate if present in the urine (Fiske, 1921a) and, in a subject who was given $NaHCO_3$, lactic acid appeared in the urine and the curves were not parallel (Brunton, 1932). Campbell (1920), Watson (1925) and Small (1928) used Leathes' method of calculation. Dodds (1923) used titration with 0.1 N. NaOH in the presence of neutral potassium oxalate to the turning point of phenolphthalein and added to the acid titration figure the acid equivalent of the "ammonia" estimated by Malfatti's formalin or Folin's aeration method. He multiplied the total by the minute volume of the urine

and so, like Henderson (1911), obtained a figure which could be correlated with synchronous gastric or respiratory investigations.

Fitz and Van Slyke (1917) preferred titration to the end point of phenolphthalein rather than to pH 7.4 because the former is closer to the reaction of the urine at which the ammonia is zero. D. W. Wilson (1923) considered that for estimates of the relative sensitivity of the two kidney mechanisms for dealing with excess acid (the alteration of the phosphates' ratio and ammonia excretion) titration to phenolphthalein was to be preferred, whereas titration to pH 7.4 with addition of the ammonia output best gave the physiological elimination of acid.

E. K. Marshall, Jr. (1922) showed that the dilute urines of diuresis or those after alkali administration lose a large quantity of carbon dioxide and change their pH within a few minutes of being voided unless precautions are taken to preserve the gas. The work of Gamble (1922) is important also in this connection. Urine has been collected under oil by Hastings, Sendroy and Robson (1925) and by Myers and Muntwyler (1928). It has been collected in an atmosphere of alveolar air by Kaye (1929). In abnormal conditions the alveolar air should obviously be provided by the subject.

Electrometric methods have been used to find the pH of the urine as voided or to titrate it to the pH of blood. The effect of the different oxidising or reducing agents likely to be found in the urine must be investigated for each method. Kaye (1929) used the quinhydrone electrode and claimed for it a probable accuracy of 0.05 pH unit. The work of Clark (1928) and the papers of Biilmann and of Le Mer (quoted by Clark) should be studied. The accuracy of the method is doubtful for urines with a higher pH than 8.0 units (La Mer and Parsons, 1923, and others) owing to decomposition and oxidation of the quinhydrone in alkaline solutions. The method is simple in use but suffers from the fact that (unless it is used for titration to the pH of blood) no mathematical expression can be obtained from the pH and the volume of the urine which can be related to other conditions. Michaelis' method of measuring the pH of urine as described by Cole (1928) is subject to the limitations that it cannot be used for urines more alkaline than 7.0 pH and does not give a titration figure which can be multiplied by the volume of urine.

In short, no mere change of urinary reaction can be correlated with such quantities as the amount of acid in the digestive secretions and so a titration figure is the method of choice for studies in the variations of urine after meals. The titration figure should be multiplied by the out-

put in unit time in order to give the acid output. D. W. Wilson's opinion quoted above on the choice between titration to pH 7.4 and 8.2 is heartily accepted by the writer. If phenolphthalein is used as a second best method because of its convenience or for comparison with previous work, the amount of free acid will be somewhat exaggerated owing to ammonia and about 4 per cent of the BH_2PO_4 will not be estimated (Chatron, 1932). If the acid-ammonia value is found (using phenolphthalein and formol) the curves obtained should be fairly parallel to those based on titration to pH 7.4 and estimation of ammonia content.

In closing this section it is only necessary to refer to ammonia production from urea as a possible source of error in analyses. As long ago as 1859 Roberts showed that the decreased titratable acidity after meals was not due to carbonates or to ammonia.

SECTION III. DIURESIS AND THE URINARY ACID. The suggestion that the "alkaline tide" is due to diuresis causing the reaction of the urine to remain closer to that of the blood could not be put forward if the conception was based on the minute acid output but has been advanced more than once when the titratable acidity or pH only has been estimated. Roberts (1859) answered it so far as the work of Bence Jones was concerned, by pointing out that Jones had reported the density of his samples and that his figures showed the absence of any great dilution. Rüdel (1892) mentioned the matter and Leathes (1919) offered it as an alternative explanation of some of Hasselbalch's results (1912). Lutz (1924) on one occasion added to a meal of 250 grams of rusks, 3 fried eggs, 30 grams of fat and 100 grams of bread as much as 3 litres of fluid without altering significantly the pH of urinary specimens taken after the meal. Meyer (1924) however found that a litre of water raised the pH from 5.20 to 5.98 in 50 minutes and to 6.22 in 1 hour 10 minutes though 300 cc. of water were almost without effect. Roberts (1859) pointed out that a given quantity of water might or might not produce a diuresis according to the need of the body for water when it was taken. Brunton and Wilson (1932) attempted to avoid diuretic effects from the water content of foods by finding first the minimum volume of water which, after a standard breakfast at 8 a.m. would produce a diuresis in each subject at noon. Driving a motor car will produce a diuresis in some subjects. On account of possible diuresis the work of some authors is open to more than one interpretation. Of his experiments with normal subjects on normal diet Kaye (1929) says (p. 190): "No food was allowed between meals but water and tobacco were allowed in moderation." Of the experiments with the stomach tube

and normal food he writes: "No food was allowed between meals and even water was discouraged" but no information is given about the composition or the quantities of fluids—soup and tea—which were taken with meals. If the volumes of urine and of gastric juice had been measured the evidence for a decreased acid output would probably have been complete, and some attempt might have been made to relate the urine more closely to the gastric secretion. As long as only change of pH or of titratable acidity is known dilution diuresis cannot be excluded as the explanation of an "alkaline tide."

SECTION IV. DIGESTIVE SECRETIONS AND URINARY ACID. Few writers on urinary acid have referred to any of the digestive secretions but the gastric juice and there is still some difference of opinion as to its relation in time to the pancreatic and intestinal juices. The writer is prejudiced in favour of the view of Pawlow (1910) that the secretion of pancreatic juice begins within a few minutes of gastric secretion in response to vagus stimulation. This juice is less alkaline than what is secreted as a result of the secretin mechanism (Still, 1931). If pancreatic juice of an alkalinity sufficient to neutralise the gastric juice be secreted almost at once after food the necessity for a decreased acid output by the kidneys vanishes. Unfortunately in human subjects this matter is not easily tested with the result that such different opinions call for attention as those which will be quoted below from Leathes (1919) and Dodds (1923).

The Philosophical Transactions of the Royal Society, London, dated 1845 contain a communication from Dr. Bence Jones in which he reports an "alkalescence" in the urine after meals. This he first attributed to "phosphate or carbonated alkali in the food" but, in an appendix to the article he says: "It seems highly probable that the quantity of acid poured into the stomach sets free alkali sufficient in some cases to make the urine alkaline." In a second article (Jones, 1849) he describes the effect on the urine of various foods and of some pure chemicals, contrasting the greater alkalescence of the urine after meat meals with its slighter alteration after vegetable meals and adding: "Dr. Prout's capital experiment on the hydrochloric acid of the stomach during digestion gives the key to these diurnal variations in the acidity of the urine and may lead to the discovery of diurnal variations in the alkalescence of the blood." It is sometimes believed that the conception of an alkaline tide originated entirely with Bence Jones but in the first of the articles the following sentences occur:

Dr. Andrews of Belfast stated to me that having observed a case otherwise in good health, in which the urine was almost invariably alkaline about two hours after breakfast, so much so as frequently to be loaded with a deposit of phosphates while still in the bladder, he was led to observe the urine of about fifteen students in good health immediately it was voided about noon. He found it alkaline in two-thirds of the cases. Whether this tendency to alkalescent urine may, as Dr. Andrews thinks, be connected with the immunity enjoyed by the inhabitants of his district from calculous affections or whether alkalescene at this period of the day is far more general everywhere than has been supposed future observations must determine. At the present time I know five physicians in whom the above phenomena at this period of the day are frequently present in a greater or less degree; and in London this alkalescence will be found in those who are considered generally healthy much oftener than is imagined.

As far as the writer knows Doctor Andrews of Belfast was forgotten by all workers on the alkaline tide from the time when his name was set down by Bence Jones until it reappeared in Dr. D. T. Davies' (1929) article. The term "alkaline tide" appears in a careful study by Roberts (1859) of the urine from a single healthy subject aged 28 years. Samples were taken hourly from 7 a.m. till 1 a.m. next day and the 1 a.m. to 7 a.m. sample covered the hours of sleep. The samples were examined for their volume, density, solids (by Christison's formula) and reaction (found by titration with caustic soda). Eleven experiments were done on normal mixed diet, seven on vegetable diet and several with purely animal food. In order to prevent overlapping effects of one meal on the next (compare Dodds, 1923) meals were only taken twice in 24 hours. Roberts confirmed Jones' finding of decreased titratable acidity in the specimens after meals. He distinguished this decrease from a "falling off in the amount of free acid which occurs after prolonged fasting" and from "a still greater falling off in hours of sleep. This second decline," he continues, "has a totally different origin and significance from the depression that follows a meal. The latter is the true *alkaline tide*." Roberts' explanation of the urinary changes during sleep is given below (section VII, p. 391). Hopkins (1898) in Schafer's *Text Book of Physiology* accepts Roberts' definition. He points out that the post-prandial alkaline tide is not a universal phenomenon and refers to factors which will influence its presence. These factors have been unrecognised by some subsequent writers and have never been thoroughly studied in a single subject at one time. Hopkins, like Quincke (1884) found periodic variations in the acidity of the urine when no food was taken. Before his article was published the "tide" had been confirmed by Maly, C. Bernard, Gley and others (Edkins, 1898; Luciani, 1913). It had been

questioned by Beneke (quoted by Roberts) who appears to have understood Jones to mean that the urine became actually alkaline after meals: he himself found occasional depressed acidity after meals although he only took 5 samples in 24 hours. Leathes (1919) could not accept gastric secretion as the explanation of the alkaline tide. In his normal and nephritic subjects the tide was usually absent after meals taken later than breakfast and occurred also in the morning when no breakfast had been eaten. The increase of acidity which he found in the urine during the night and the fall of alveolar carbon dioxide tension which he found when subjects awoke led him to the conclusion that

> the traditional ascription of the alkaline tide to the secretion of hydrochloric acid in the stomach appears to be erroneous. An intense alkaline tide is normal in the morning when no food is taken; alkaline tides after other meals than the first of the day are certainly not the rule and in the experience of the writer do not occur. It is not even *a priori* probable that they should, since the passage of acid from the stomach is the signal for the secretion of alkaline digestive juice into the small intestine and for absorption into the blood to begin. . . . The experiments gave support to the hypothesis that the morning alkaline tide is related to variation in the activity of the respiratory centre. (This relationship will be discussed in section vi, p. 388, of this paper.)

Campbell (1920), investigating Europeans and Eurasians at Singapore, found an alkaline tide (in Leathes' sense) after all meals and felt that his results supported those of Hasselbalch (1912) rather than those of Leathes. Campbell and Webster (1921) obtained a good alkaline tide in more than half the midday meals taken by a single subject. Dodds (1923) differing from others recognised, as Henderson (1911) had done, the importance of the ammonia excretion as well as the urinary volumes. On the basis of the acid-ammonia excretion in unit time he found an "alkaline tide" after the first meal of the day which was followed during the fourth and fifth hours approximately by an acid tide. He held that this acid tide masked the appearance of alkaline tides after subsequent meals. This explained Leathes' results but not those of Hasselbalch or Campbell. Dodds' previous work on the alveolar CO_2 and on the CO_2 of the blood after meals (Dodds, 1921; Dodds and Bennett, 1921; Dodds and McIntosh, 1923; Dodds and Smith, 1923) led him to believe that, during the first 30 to 45 minutes after a meal the production of gastric juice was accompanied by a compensatory retention of carbon dioxide in the blood, while between 1 and 2 hours after the meal the maximum pancreatic secretion was accompanied by a fall of blood carbon dioxide below the subject's normal level. It might have been expected, then,

that his subjects would have had their "alkaline tide" during the first hour after food but in most of the experiments the "tide" was later. It is doubtful whether this can be explained by a lag during which the alkali that remained in or was returned to the blood (owing to the production of gastric juice) would at first be neutralised by retained CO_2 and later would be gradually excreted by the kidneys. The difficulty arises from the fact that the alveolar CO_2 does not always rise when a very large secretion of gastric juice takes place (Brunton and Israëls, 1930). Further, the kidney responds very quickly to increase in alkaline substances such as $NaHCO_3$ which may be introduced into the circulation through the digestive tract (Henderson and Palmer, 1913b). The explanation of the rise of alveolar CO_2 by gastric secretion has been also criticised on mathematical grounds by Brunton and Israëls (1930). On the other hand, Dodds' method of calculating the changes of urinary acid excretion seems the most fruitful method available. Dodds himself reports, with scientific candour that "observations on a gastrectomised man were so irregular as to be worthless and are not reproduced;" also that "it was hoped that by examining men who secreted varying quantities of hydrochloric acid, urinary changes would be found. Unfortunately this was not found to be the case." Would that all investigators were equally willing to offer their difficulties as signposts for those who seek to follow in their steps! Other investigators who have studied the urine or alveolar CO_2 of achlorhydriacs include Hubbard, Munford and Allen (1924), Watson (1925), Munford and Hubbard (1926), and Hubbard and others (1931). Dodds abolished the decrease in acid-ammonia output after a meal by the use of atropin and emphasised the need for a suitable test meal to eliminate achlorhydriacs before expecting a decreased acid-ammonia output after meals. Schulten (1928) and Kaye (1929), on the other hand, point out the need for an efficient gastric stimulus to avoid failure in producing a rise of urinary pH or a false diagnosis of achlorhydria. Schulten found that the absence of gastric secretion after a mixed meal went with a failure to increase in the pH of the urine. The greater titratable alkalinity found to follow protein meals (Jones, 1849; Kaye, 1929; Hubbard and Steele, 1929) and discussed below in section Va, page 385, have been thought to support the theory of the gastric origin of the alkaline tide since proteins are good stimuli for gastric juice. D. T. Davies (1929) working in Dodds' laboratory reports on the changes of urinary pH in normal and abnormal subjects who were given the standard gruel test meal. Specimens were taken half hourly and examined with the hydrogen electrode as soon as pos-

sible. Stomach contents were withdrawn regularly for several hours, the patients being either in bed or resting during the time. The results and curves published show that the gastric secretion and a rise of urinary pH occurred or were absent together. An acid tide was obtained in an achlorhydriac whose stomach emptied so as to be free of starch in ¾ hour. The "alkaline tide" was absent in a subject who salivated 400 cc. of saliva. In the majority of 20 subjects who had free hydrochloric acid the pH of the urine rose one hour after the meal and remained raised during one or more half-hour periods. The writer says: "Many fallacies and sources of error exist but if the estimations are done within a short time of voiding the urine and especially if repetition of the test meal produces no change in the urine or no alkaline tide then it can reasonably be assumed that this is a case of absolute achlorhydria." Davies may be said to have confirmed the existence of a raised pH due to meals and, if his conclusion is confirmed regarding the diagnostic power of the urinary examination, to have furnished a very useful clinical test to those who can undertake a satisfactory pH examination. The volumes of urine passed, the titratable acidity and the ammonia content were not reported nor was the question discussed of an "alkaline tide" before the morning meal.

Brunton and Wilson (1932) claim to have produced good gastric secretion without an alkaline tide by meals of bovril and biscuits at 1:00 p.m. On one occasion the free acid of the gastric contents during consecutive intervals of 15 minutes beginning 30 minutes after the commencement of the meal was 12 cc., 39 cc., 73 cc., 73 cc. and 24 cc. of 0.1 N. acid per 100 cc. of contents. Yet no "alkaline tide" took place, in the sense of decreased minute acid output. The failure might be explained on Dodds' hypothesis that it was inhibited by an acid tide due to breakfast. Later however a meat meal was given to a subject at 9:15 a.m. 12 hours after his previous meal. At 10:35 his stomach contents were withdrawn and showed free acid by Gunzberg's and the congo red tests; on titration to Töpfer's reagent the "free acidity" was 72.0 cc. and to phenolphthalein 88.0 cc. of 0.1 N. acid per 100 cc. of contents. In spite of these figures the acid output (and titratable acidity) of the urine increased steadily from 9:15 a.m. till 1:00 p.m. The acid-ammonia output ran parallel with one slight exception: in the 11:00 to 11:30 specimen it fell from 0.75 to 0.725 cc. of 0.1 N. acid per minute, rising again to 0.745 in the next period. At least ⅓ of the meal was present in the stomach 80 minutes after the beginning of the meal so it is doubtful whether the result can be explained by very rapid empty-

ing of the stomach as was one of Davies' (1929) pH results. If such observations are confirmed it will be clear that gastric secretion can occur without producing a decreased acid output of urine. The allied question of whether decreased urinary acid output can occur in the absence of gastric secretion is answered affirmatively in sections VI and VII. That gastric secretion is accompanied by an increased pH of the urine seems proved by the work of Davies described above, but their simultaneous occurrence does not prove the dependence of the one on the other. Still less has any numerical relationship between them been established.

As regards the other digestive secretions Campbell and Webster (1921) suggested salivary secretion as the cause of a transient increase in the acidity of the urine which preceded the "alkaline tide" after certain meals. Davies (1929) noted a case where the subject produced 400 cc. of saliva after a test meal and, although he had "free hydrochloric acid" in the stomach, had no rise of pH in the urine. Dodds (1923) considered that the alkaline tide was followed by an acid tide lasting about 3 hours and due to pancreatic secretion. This acid tide would mask the alkaline tide after a second meal such as lunch or supper. An alternative explanation suggested by the work of Kaye (1929) has been mentioned and other possible factors will be considered in section 5.

SECTION V. FOOD CONSTITUENTS AND URINARY ACID. (*a*) *The stimulating effect of food on gastric secretion.* There is an almost universal opinion that protein meals are the best for the production of an alkaline tide (Kaye, 1929). They have been contrasted in this respect with carbohydrate meals by Bence Jones (1849), Roberts (1859), Hopkins (1898), Howe and Hawk (1914) Robinson, Shearns and Davies (1924), Hubbard, Munford and Allen (1924), Meyer (1924), Kaye (1929) and Hubbard and Steele (1929). Most of the meals have been taken at breakfast or lunch time. Lutz (1924) used mixed meals of vegetables and meat which he found better than eggs. Hubbard and Steele reported urinary conditions after breakfasts of different composition. The alkaline tide (as judged by pH) following a normal breakfast with steak added was greater and began sooner (both when the subject took it in bed and when he rose first) than a breakfast without a steak. Kaye's table 5 and curve 8 (high protein diet with hourly samples for 24 hours) compared with his tables 6 and 7 and curve 9 (carbohydrate diet) show clearly that, judged by pH alone, a protein diet produced more strongly alkaline tides after all meals in his subjects. On the day shown in curve 8, identical meals of 12 ounces of cooked meat only were taken at 8 a.m., 1 p.m. and 6 p.m. On four days when the diet consisted

of bread, milk and honey the morning tides were well marked but no tides followed the midday meals. A meal at 4 a.m. consisting of bovril was followed by gradually increasing urinary pH in the specimens collected at 6, 7 and 8 a.m. No such increase occurred after either of two meals consisting of 80 grams of barley sugar at 3 a.m. Fiske (1921a) obtained a rise of pH with meat where no rise occurred with carbohydrate foods. Lutz (1924) reported the absence of a tide after an oatmeal meal, while Meyer (1924) reported its absence after olive oil, glucose and levulose. The fact that these substances are poor gastric stimuli would reinforce the evidence for a gastric explanation of decreased urinary acid after meals if similar results are obtained when the total acid output is measured. Rannenburg (1925-26) found two elevations in the urinary pH during the day: in the morning and afternoon. These, however, disappeared if a meat or meat and fat diet was continued for several days: a result which she could not explain. Small (1928) added haddock, cheese, ham, banana and apple to a standard breakfast of a roll and cup of tea. With Leathes' "percentage alkalinity" as her criterion, she found a marked alkaline tide after haddock and a definite one after banana but none after the other foods. She argued that if the tide after haddock had been due to protein, the haddock (12 per cent), the cheese (26 per cent) and the ham (16 per cent) should have produced tides of the same order. As regards the fat content of the two latter, she measured the effect on the tide of adding from $\frac{1}{4}$ to $1\frac{1}{4}$ ounces of butter to the roll and tea. The effects of the different amounts could not be explained. She held that the banana and apple had approximately the same excess of inorganic basic radicles so that the difference of their effects on the urine did not depend on that factor alone. She did not measure the volumes of the urine or consider the urine passed after waking and before 8 a.m. Kaye (1929) published curves based on hourly analyses of gastric contents and urine. The stimulus to the stomach was a mixed meal the nature of which is fairly fully described. The gastric contents (curve 12) had about the same pH 3 hours after breakfast and lunch but a decidedly greater titratable acidity after breakfast. The pH of the urine was about the same at both these times. Five meals consisting of meat, or eggs, with toast, tea and a sweet taken at 6 p.m. failed to raise the pH of the urine (table 3 and curve 6). Two meals at the same hour with gastric analyses showed a high initial gastric acidity and produced no further rise of gastric acidity nor any alkalinity of the urine: a result explicable on the hypothesis that the stomach was secreting to its maximum power before the meal

(if it is assumed that the minute volume of gastric juice remains unaltered). With the work of Maclean and Griffiths (1928) in mind, Kaye gave hydrochloric acid by mouth and abolished the rise of urinary pH after meals. Yet it is possible that the abolition was due to ingested hydrochloric acid as much as to the inhibition of secretion in the stomach.

Gastric secretion has been thought to produce a rise of urinary pH. Naturally this tide would be increased by foods which are good gastric stimulants. As, however, the rise of pH may occur without much fall in the titratable acidity of the urine (Fiske, 1921) and since the minute output of acid does not always run parallel to the conditions in the stomach (Dodds, 1923; Brunton and Wilson, 1932) the effect of food constituents on the acid or acid-ammonia output probably depends on other factors beside the gastric secretion:—the acid-base balance of the food constituents absorbed (see below, section Vb), metabolic conditions in the body and the respiratory mechanism (see section VI).

Food constituents and urinary acid. *b. Composition of foods.* Apart from any stimulant action on the gastric juice, foods might bring about decreased acid output in several ways. Hasselbalch (1912) suggested that carbohydrate food might decrease the use of protein in the body and so the output of acid would be diminished after food. The discovery of the rise of alveolar CO_2 after meals by Higgins (1914) and Erdt (1915) made this explanation less acceptable and it was criticised also by Fiske (1921a) on the grounds, first, that before his "alkaline tide" disappeared Hasselbalch's subject had been without carbohydrate food for 4 days and might have been suffering from acidosis; secondly, that estimations of urinary sulphates show that the sulphate output decreases during the day and does not rise at the time when the urine is acid; and lastly, that the decrease in the sulphate is accompanied by a parallel decrease in the ammonia output. Fiske (1921b) also showed that the organic acids of the urine (estimated by Van Slyke and Palmer's method) were constant and so were not the cause of increased pH after meals. Lutz (1924) Kaye (1929) and others failed to produce rises of urinary pH with carbohydrate meals.

A number of workers on total metabolism have studied the effect of different foods and especially of their ash on the 24 hour specimen of urine but this approach to the problem of the "alkaline tide" has only recently been attempted. Bunge (1890), for instance, taught that the large potassium content of potato displaces the sodium and forms a compound with the chlorine of the blood, the sodium then uniting with H_2CO_3 to form bicarbonate. These substances being in excess are ex-

creted by the kidney so that the body may actually suffer from serious loss of NaCl, as well as excrete an alkaline urine. Sherman and Gettler (1912) contrasted the neutral S, the total N, the NH_3 and the titratable acidity of the urine in 3 periods of 3 days each when a diet with more acid ash was substituted for the previous one. They found that the extra acid of the second diet was excreted as regards $\frac{1}{3}$ by extra ammonia output and as regards $\frac{2}{5}$ by extra titratable acidity, that is, by a larger ratio of BH_2PO_4/B_2HPO_4. In another series of meals the substitution of rice for potato was followed by a rise in the acidity of the 24 hour specimens. (As Lusk, 1928, points out, their researches, with those of Hindhede and others, should raise the potato to a high level of dietary dignity.) C. Wilson and the writer (1932) produced a large decrease in the acid output of the urine by giving a midday meal of potato and smaller but quite definite decrease with banana (compare Small, 1928). By superimposing potato on a meat meal they produced a larger decrease of acid output than when the meat was given alone: a finding contrary to what would be expected if carbohydrates did not produce an "alkaline tide" in their sense. They also studied a number of other foods separately. Taylor (1932) has produced an alkaline tide in an achlorhydriac patient by using a potato meal. If the presence of a tide in Davies' (1929) test is to be used for elimination of achlorhydria his standard meal must be adhered to and mixed meals of unknown composition must not be given. Blatherwick and Long (1922) added orange juice to a standard diet and found that not only was the pH of the 24 hour sample raised but that the CO_2-combining power of the blood rose one hour after taking 800 cc. of juice from 63.3 to 67.2 vol. per cent. It is reasonable to suppose that the citric acid of the oranges disappeared in the body and that the alkali radicles liberated raised the CO_2-combining power of the blood until they were excreted by the kidney (citric acid may be converted to glucose in the body).

The composition of foods as regards their inorganic radicles and their content of organic acids which do not undergo changes in the body must be considered before a complete explanation of urinary changes after meals can be attempted.

Section VI. The respiratory centre and urinary acid. A change in the excitability of the respiratory centre was suggested by Leathes (1919) in order to explain the occurrence of increased "percentage alkalinity" of the urine which he found in the morning whether a meal was taken or not. It is an hypothesis advanced to explain the alkaline tide after sleep (the subject of the next section), and is quite

distinct from the effect on the respiratory centre produced by a loss of hydrochloric acid from the blood into the stomach during the production of gastric juice. During the latter activity the body might correct its loss of hydrions either by retaining more H_2CO_3, by excreting more alkali through the kidneys or by both methods. It is quite possible that different subjects respond differently to the same circumstances. If the respiratory centre receives less stimulation owing to the lowered plasma CO_2, ventilation will be decreased, CO_2 be retained and the alveolar CO_2 will rise. Dodds and his co-workers undertook the investigation of this hypothesis. The alveolar CO_2 tension was found to rise by about 2 to 6 mm. Hg between $\frac{1}{2}$ and $\frac{3}{4}$ hour after meals and, having returned to normal about $\frac{1}{2}$ hour later, to fall below the normal limit for one or two half-hour periods (Dodds, 1921; Dodds and Bennett, 1921; Bennett and Dodds, 1921). The carbon dioxide content of the blood was found to rise (Dodds and McIntosh, 1923) and the plasma chlorides to fall (though the corpuscular chlorides rose) after meals (Dodds and Smith, 1923). The changes in the urine have already been described and some difficulties pointed out (section IV, p. 380) but the work has been generally accepted. (Wilson, 1923, and others) as explaining the change in urine after meals. Van Slyke, Stillman and Cullen (1917) argued that if the respiratory centre was less sensitive than normal after meals the ratio of the alveolar CO_2 to the CO_2-combining power of the plasma should alter. In four out of thirteen experiments the ratio Plasma CO_2/Alveolar CO_2 changed in value by less than 0.02 which was the limit in error. In the other nine experiments it was definitely increased. The plasma bicarbonate sometimes rose slightly after meals and sometimes did not. The authors concluded that more tests were needed but Peters and Van Slyke (1931, p. 965) report the experiments as showing that "the rise of alveolar carbon dioxide was accompanied by a similar change of serum bicarbonate." Hermann and Saker (1927) studied the alkali reserve (CO_2-combining power) of the plasma before and after meals in seven hyperchlorhydriacs and in ten hypochlorhydriacs (tested by histamine). The resting level of the former was on the lower limit of the normal and its rise after food was greater than in the normals studied by Van Slyke, Stillman and Cullen. If those hypochlorhydriacs who gave a delayed secretion are omitted from consideration, the others with one exception had a resting alkali reserve at the upper normal limit. Their reserve was little affected, unaltered or even lowered by the breakfast.

Gastric secretion is, however, not the only factor which will influence

the respiratory centre. Leathes (1919) produced an increased percentage alkalinity of the urine by voluntary hyperpnea. Hyperpnea due to rise of body temperature at rest might be expected to do the same. Experiments by Talbert (1919) show, however, that when subjects sit in a hot cabinet at 40° to 42°C. for 15 minutes their urinary pH at all events tends to fall. (A similar fall was found and was to be expected when the rise of temperature was produced by exercise.) But since the volumes of urine were not reported and are likely to fall during sweating it is impossible to say whether the acid output was increased or decreased. In a further series of experiments (Talbert and others, 1931) the blood pH was found to rise very slightly if at all beyond the limits of error while the CO_2-combining power of the plasma usually fell about 3 to 5 volumes per 100 cc. of blood.

Besides the effects on the respiratory centre of gastric secretion, hyperpnea and raised temperature, the effect of posture has been considered. Higgins (1914) found that when subjects lay down the alveolar CO_2 rose above the level at which it was when they were standing or sitting upright. He suggested tentatively that the connection between this rise and the rise which he also observed after meals might be found in vasodilatation in the splanchnic area. He considered that coffee would produce vasoconstriction and, on testing it, obtained a fall of alveolar CO_2. (Simpson, 1929, has found changes in the output of urinary phosphate on lying down.) (See below.)

The changes of the blood gases which accompany sleeping and waking have been studied by Leathes (1919), Collip (1920), Collip and Backus (1920), Bass and Herr (1922), Endres (1923), Watson (1925) and others.

Leathes found that if subjects were awakened during the night the alveolar CO_2 was higher than before going to bed or than after getting up. Further, the alveolar CO_2 on waking was about 0.7 cc. per 100 cc. higher than after getting up. Collip found either no alteration or a fall in the CO_2-combining power of the plasma when students were awakened at different times during the night. He concluded than an actual increase of pH took place in the blood during sleep. Collip and Backus found that the results of 15 to 37 minutes' hyperpnea included a fall of alveolar CO_2, a fall of CO_2-combining power in the plasma, a marked decrease of urinary acidity and ammonia excretion with an increase of urinary phosphate concentration and elimination. Bass and Herr confirmed Leathes' findings and extended them by observing a maximum CO_2 tension at about 45 minutes after the onset of sleep. If the subject was wakened subsequent sleep was accompanied by a new

rise of CO_2 in the alveolar air. In six experiments Endres (1923) took samples of blood for CO_2-combining power estimations at the instant when subjects were awakened and compared them with the alveolar CO_2 tensions obtained immediately afterwards. He confirmed the increase of CO_2 tension and found the CO_2-combining power unchanged. Similar results have been obtained by Gallwitzer-Meier and Kroetz (1924).

Besides confirming the results of Leathes, Watson obtained increases of CO_2 both at 9 and at 11 a.m. in a subject who had breakfast at 8:45 a.m. He obtained an increased "percentage alkalinity" in the urine of achlorhydriacs and in a subject to whom atropin had been given in amount proved by analysis to have prevented gastric secretion. He concluded, as Roberts had done, that the morning tide was separate from the tide after meals.

The respiratory centre, then, may be affected during gastric secretion, during voluntary hyperpnea and possibly during sweating or change of posture. It appears to be less sensitive during sleep.

Section VII. Sleep, waking and urinary acid. The problem of sleep has recently been the subject of a review by Kleitman (1929) on which table 1 is largely based.

It will probably be agreed that, however interesting are the changes produced by sleep in the phosphate and chloride excretion and the change produced in the former by posture, no complete account of the metabolism of sleep can be offered until further work is done. Among factors which must be considered before the tale is fully told are 1, fasting; 2, products of food eaten during the preceding day; 3, posture; 4, alteration of excretion by sweat; 5, altered balance of internal and digestive secretions; 6, altered activity of the kidney; 7, altered metabolism due to muscular relaxation; 8, alteration of sensory stimuli due to quiet, relaxation closing of eyes, etc., and especially altered temperature of body surroundings while in bed; (9) inhibitions from higher centres as suggested by the work of Pavlov (1929) and of Hess (1929).

Fasting has been offered as an explanation of the rise of acidity during the night and its fall after breakfast (Hasselbalch, 1912; Kaye, 1929) but will not explain the morning decreased acid output unless a gastric secretion is believed to occur in the absence of breakfast,—a belief which the writer believes untenable for reasons given above. Posture has also been considered in section VI, page 388. Depression of the respiratory centre (Leathes, 1919) might be due to any of the factors 4, 7 or 8, singly or together.

TABLE 1

Changes in the urine and venous blood during sleep, changes of posture, and different periods of the day

	Urine	
Volume in unit time	Fell during sleep	Kleitman (1923); Simpson (1924)
pH	Fell during night	Roberts (1859); Leathes (1919)
	Fell to about pH 5.3 during sleep	Kaye (1929)
Acid output	Decreased during sleep	Roberts (1859)
	Increased during sleep, night or day	Campbell and Webster (1921–22)
	Increased during sleep, night or day	Kleitman (1923)
P_2O_5 output	Usually not altered	Simpson (1926)
	After a day's fast, lower in sleep	Simpson (1929)
	After a day with only 2 meals, higher	Simpson (1929)
	After a day's fast and subject up all night, output higher	Simpson (1929)
	Higher when lying down than standing	Simpson (1929)
	Higher during sleep	Campbell and Webster (1921)
	Higher during sleep	Kroetz (1926)
	Higher during sleep whether night or day	Kleitman (1923)
	Max. output 4–7 p.m. or 7-midnight	Cathcart, Kennaway and Leathes (1908)
	Max. output at night, minimum forenoon	Wordell (1924)
	Fell when standing active	Haldane, Wigglesworth and Woodrow (1924)
	Fell after dinner, if not prevented by voluntary hyperpnea	Broadhurst and Leathes (1920)
Cl. output	Fell at night, rapid increase in morning	Norn, cited by Peters and Van Slyke (1931)
	Fell during sleep, night or day	Kleitman (1923); Simpson (1926)
NH_3 content	Increased during sleep	Kroetz (1926)
NH_3 output	About twice as great during sleep	Fontés and Yovanovitch (1923)
Uric A. output	Low at night: Max. in morning	Cathcart, Kennaway and Leathes (1908); Campbell and Webster (1921)

TABLE 1—*Concluded*

	Urine—Concluded	
Ca output	Decreases during sleep	Kleitman (1923)
SO_4 output	Highest during periods of alk. tide	Fiske (1921a)
	Increased by activity	Campbell and Webster (1921)
	Blood	
P_2O_5	Raised during sleep	Haldane, Wigglesworth and Woodrow (1924)
Na and Cl. and P_2O_5	Raised 2 hours after onset of sleep	Gallwitzer-Meier and Kroetz (1924)
HCO_3, K and Ca	Unchanged 2 hours after onset of sleep	Gallwitzer-Meier and Kroetz (1924)
Ca	Increased during sleep in 5 psychopaths	Heilig and Hoff (1925)
	Falls during sleep by 6 to 8 per cent	Demole (1927)
CO_2-combining power	Equal or decreased during sleep	Collip (1920)

Roberts (1859) believed that the falling off in the hourly excretion of acid during sleep was due "simply to the diminished secretion of solid urine (by which he meant urinary solids) which . . . . invariably accompanies the state of fasting and sleep. That this is the true interpretation of the fact," he continues, "may be seen by comparing the amount of free acid corresponding to 100 grams of solid urine with the amount separated per hour. It is then seen that the *degree of acidity* of the urine steadily rises after recovering from the depression of the alkaline tide until the next meal." He adds that there is a limit to the increase of acidity during a prolonged fast since an extra 12 hours' fast did not increase the acidity of urine in experiments of Bence Jones. Roberts' figures are given in table 2 for urinary acidity and for "measures" of acid per hour, each "measure" containing $1\frac{1}{2}$ of a grain of sodium carbonate. The findings of Leathes (1919), Collip (1920) and Watson (1925) regarding the acidity of the urine, the carbon dioxide content of the alveolar air and the CO_2-combining power of the blood during sleep have been described in the preceding sections.

Kaye (1929) measured the pH of two-hourly or hourly samples of urine during the night from a subject who was awake and performing work in a warmed laboratory, in order to see if the wakefulness would prevent any decreased sensitivity of the respiratory centre and so would eliminate the morning rise of urinary pH. The pH continued to fall

throughout the night and until after breakfast which was taken at 8 o'clock: there was no "alkaline tide" (curves 3 and 4). When bovril was taken at 3 a.m. the pH of about 5.3 rose to 6.0 in the 6 a.m. sample, and to 6.8 in the samples at 7 and 8 a.m. On two occasions meals of 80 grams of barley sugar at 4 a.m. were not followed by a rise of pH though in these experiments the subject was in bed (table 6 and curve 9). Kaye writes (p. 194): "the present paper seems to show definitely that the nocturnal acidity is due to starvation since it may be abolished by taking (non-carbohydrate) food." He explains the failure of carbohydrate food to produce a rise of pH by the fact that it is a poor stimulant to gastric HCl-secretion (p. 202) and without relation to starvation conditions. On eight occasions breakfast was omitted and no food taken till 1 p.m. An "alkaline tide" was found on three occasions: one curve was of an anomalous flat type, and on four occasions there was no al-

TABLE 2

Changes in urinary acidity and acid output during the night

Roberts 1859

TIME	ACIDITY IN GRAINS PER 1000 GRAINS OF URINE	ACIDITY PER HOUR IN "MEASURES" (1 MEASURE = $\frac{1}{12}$ GRAIN)	ACID PER 1000 GRAINS OF URINARY SOLID
7–9 p.m.	0.55	0.48	1.02
9–11	0.93	0.77	2.02
11–1 a.m.	1.07	0.62	2.13
1–7	1.30	0.38	2.37
7–8	1.55	0.45	2.50

kaline tide. Kaye does not state the condition of the urine previous to 8 a.m. in this series of experiments or give any information about the volumes of urine passed. Hubbard (1930) has published results based on examinations of about 370 urines hourly from 7 a.m. onwards. The subjects were wakened at 7 a.m. and had breakfast between 8 and 9 o'clock. Some may have dozed again between 7 and 8 o'clock. Frequency curves were constructed for the pH of the urines at 7, 8 and 9 a.m. which showed a rise of the average and median pH's in the period from 7 to 8 o'clock. After referring to cases of achlorhydria (Hubbard and Munford, 1926) which showed "alkaline tides" Hubbard says that "the explanation of the rise in pH is probably the one offered by Leathes: immediately after waking accumulated carbon dioxide is blown off and an alkalinity of the urine results just as it does in similar experiments carried on at any time." In practically every case where there was a further rise of

pH at 9 a.m. there was hydrochloric acid in the gastric juice. Hubbard has thus come to distinguish two causes of increased urinary pH:—gastric secretion and respiratory hyperpnea. His curve of the differences between the 8 and 9 a.m. samples, plotted on probability paper, is not a true distribution curve but shows that various factors influence that difference. Some possible factors have been suggested in the present review. With Hubbard's article, Roberts' assertion of the difference between the true alkaline tide and the morning tide may be said to be generally accepted.

It still remains to prove the exact changes in the respiratory centre. Why should it produce such a degree of ventilation when the subject wakes as to excrete too much CO_2 so that the kidney has to correct the balance? The respiratory centre has, as a rule, a very tender conscience. What makes that conscience over-active at the waking hour? Is the centre stimulated then in a special way apart from the raised CO_2 content of the blood? How is the phenomenon related, if at all, to the changes in the blood and urine chemistry mentioned in table 1? When we can answer these questions we shall know more than we do now of the nature of sleeping and waking.

Hubbard and Steele (1929) were the first to record experiments in which an effort was made to distinguish between waking and rising in the production of the morning fasting "alkaline tide." One of their subjects was roused at 7 a.m. and specimens were taken hourly for examination of pH by indicator methods. If no meal was given the subject voided a urine of higher pH when in bed than when he rose; there was an increased urinary pH when he drank water in bed but not when he got up first; there was a "tide" when he ate a normal breakfast in bed but it was delayed (on two occasions) when the meal was eaten after rising; lastly tides of about equal alteration of pH and beginning about the same time followed a "normal" breakfast to which a large portion of steak was added, whether eaten before or after rising. Taking as his criterion of a "tide" changes in the minute acid-ammonia output the writer (Brunton, 1932) studied two subjects neither of whom took any food or fluid on waking. The first showed a decreased acid-ammonia output once when he remained in bed and one decreased and one increased output when he rose as soon as he wakened. Judged by the titratable acidity of the urine plus its ammonia content per 100 cc. there was an alkaline tide when he stayed in bed and an acid tide (twice) when he rose on waking. The second subject remained in bed on five occasions after he woke. In each case he had a decreased acid-ammonia output

during the first half hour. On two of these occasions he remained in bed for a second half hour during which the decrease was still greater. On the three occasions when he rose after the first half hour the minute acid output continued to fall while the titratable acidity continued to fall in one and rose in two. On three different occasions when the subject rose on waking, the minute acid output decreased during the first half hour while the titratable acidity rose on one occasion and fell twice. On another occasion he woke at 3:55 a.m. and stayed awake until 4:25 a.m. The urine passed during this period showed a decreased acid-ammonia output compared to the urine passed between 11 p.m. and 3:55 a.m. (suggesting confirmation of Roberts, 1859). He then slept, being awakened at 5:05 a.m. and again at 5:45 a.m. During these two periods both the minute acid excretion and the minute acid-ammonia excretion rose (though the titratable acidity continued to fall). At 6 a.m. he got up and the urine from 5:45 to 6:20 a.m. again showed decreased acid and acid-ammonia outputs while the titratable acid tended to rise. (This experiment is detailed not only to suggest the connection between sleep and the urine but also to emphasize the different results obtained according to whether or not the volumes of urine are considered.) If the suggestion of these three workers' experiments is borne out by further work, the decreased acid output of the morning will be explained by the act of waking rather than by rising from bed. Further confirmation is required and even further evidence before a complete explanation of the "tide" can be offered.

SUMMARY

1. The conception of the alkaline tide of the urine in the original sense of decreased titratable acidity or the later sense of increased pH is contrasted with the conception of increased acid-ammonia output per minute (or hour). The latter conception allows a numerical relation to be sought between the acid of the urine and other body changes: the former is inferior in this respect.

2. Reference is made to various methods of measuring the urinary acid.

3. Digestive activity may be accompanied by a deceased acid output in the urine. This has usually been held to depend on the loss of acid by the blood during the production of gastric juice.

4. Food may reduce the urinary acid output either through excess of its basic radicles, through preferential absorption in time of its alkali constituents, by the oxidation in the body of its organic acid or by saving protein metabolism.

5. Voluntary hyperpnea can produce a fall of CO_2 in the blood and a decreased acid output in the urine. The latter is accompanied during sleep by a rise in the output of phosphates and sometimes by a fall in the CO_2-combining power of the blood. Evidence is not available for a complete description of the phenomena. A fall of blood CO_2 and a decreased acid output occur immediately on waking. Their cause is not completely understood but they are certainly not due to food and probably not due to getting out of bed.

BIBLIOGRAPHY

Bass and Herr. 1922. Zeitschr. f. Biol., lxxv, 279.
Bennett and Dodds. 1921. Brit. Journ. Exp. Path., ii, 58.
Blatherwick. 1914. Arch. Int. Med., xiv, 409.
Blatherwick and Long. 1922. Journ. Biol. Chem., liii, 103.
Broadhurst and Leathes. 1920. Journ. Physiol., liv, Proc. xxviii.
Brunton. 1933. Journ. Physiol., lxxviii, 65.
Brunton and Israëls. 1930. Journ. Physiol., lxx, 184.
Brunton and Wilson. 1932. Journ. Physiol., liv, 10 P.
Bunge. 1890. English transl. by Wooldridge, Text book of physiology and pathological chemistry.
Calvert, Mayrs and Milroy. 1921. Journ. Path. and Bact., xxiv, 91.
Campbell. 1920. Biochem. Journ., xiv, 603.
Campbell and Webster. 1921. Biochem. Journ., xv, 660.
1922. Biochem. Journ., xvi, 507.
Cathcart, Kennaway and Leathes. 1908. Quart. Journ. Med., i, 406.
Chatron, M. 1932. Bull. d. l. Soc. d. Chim. Biol., xiv, 783.
Clark. 1928. Determination of hydrogen ions. Baltimore and London.
Cohen and Dodds. 1923. Journ. Physiol., lix, 259.
Cole. 1928. Practical physiological chemistry. Cambridge, p. 60.
Collip. 1920. Journ. Biol. Chem., xli, 473.
Collip and Backus. 1920. Amer. Journ. Physiol., li, 568.
Davies. 1929. Brit. Journ. Exp. Path., x, 1.
Demole. 1927. Arch. Exper. Path. and Pharm., cxx, 229.
Dodds. 1921. Journ. Physiol., liv, 342.
1923. Brit. Journ. Exp. Path., iv, 13.
Dodds and Bennett. 1921. Journ. Physiol., lv, 381.
Dodds and McIntosh. 1923. Journ. Physiol., lvii, 139.
Dodds and Smith. 1923. Journ. Physiol., lviii, 157.
Edkins. 1898. In Schafer's Textbook of physiology, vol. 1.
Endres. 1922. Biochem. Zeitschr., cxxxii, 220.
1923. Biochem. Zeitschr., cxlii, 53.
Erdt. 1915. Deutsch. Arch. klin. Med., cxvii, 479.
Fiske. 1921. Journ. Biol. Chem., xlix, 163.
1921b. Journ. Biol. Chem., xlix, 171.
Fitz and Van Slyke. 1917. Journ. Biol. Chem., xxx, 389.
Folin. 1903. Amer. Journ. Physiol., ix, 25.

Fontés and Yovanovitch. 1923. C. R. Soc. Biol., lxxxviii, 456.
Gallwitzer-Meier and Kroetz. 1924. Biochem. Zeitschr., cliv, 82.
Gamble. 1922. Journ. Biol. Chem., li, 295.
Gamble, Ross and Tisdall. 1923. Journ. Biol. Chem., lvii, 633.
Goiffon, R. 1931. C. R. Soc. Biol., cvi, 174, 808, 904, 1223.
Haldane, Wigglesworth and Woodrow. 1924. Proc. Roy. Soc. London B. xcvi, 1.
Hasselbalch. 1912. Biochem. Zeitschr., xlvi, 416.
Hastings, Sendroy and Robson. 1925. Journ. Biol. Chem., lxv, 381.
Heilig and Hoff. 1925. Quoted by Kleitman, 1931.
Henderson. 1911. Journ. Biol. Chem., ix, 409.
Henderson and Palmer. 1913a. Biol. Chem., xiii, 393.
1913b. Biol. Chem., xiv, 81.
1914. Journ. Biol. Chem., xvii, 306.
Herman and Saker. 1927. Klin. Wochenschr., vi, 1367.
Hess. 1929. Abst. Communications 13th International Physiol. Congress, p. 119.
Higgins. 1914. Amer. Journ. Physiol., xxxiv, 114.
Hopkins. 1898. In Schafer's Text book of physiology, vol. 1.
Howe and Hawk. 1914. Journ. Biol. Chem., xvii, xlviii.
Hubbard. 1929. Biol. Chem., lxxxiv, 191.
1930. Amer. Journ. Physiol., xci, 618.
Hubbard, Munford and Allen. 1924. Amer. Journ. Physiol., lxviii, 207.
Hubbard and Steele. 1929. Journ. Biol. Chem., lxxxiv, 199.
Hubbard et al. 1931. Journ. Biol. Chem., xcii, Proc. Soc. Biol. Chem., xxix.
Jones. 1845. Phil. Trans. Roy. Soc. London, cxxxv, 335.
1849. Ibid., cxxxix, 235, 261.
Kaye. 1929. Austr. Journ. Exp. Biol. and Med. Sci., vi, 187.
Kleitman. 1923. Amer. Journ. Physiol., lxvi, 67.
1929. Physiol. Rev., ix, 624.
Kroetz. 1926. Zeitschr. ges. exp. Med., lii, 770.
La Mer and Parsons. 1923. Journ. Biol. Chem., lvii, 613.
Leathes. 1919. Brit. Med. Journ., ii, 165.
Luciani. 1913. Transl. Welby, Text book of physiology, ii, 379.
Lusk. 1928. The science of nutrition, p. 485.
Lutz. 1924. Zeitschr. ges. exp. Med., xli, 516.
Maclean and Griffiths. 1928. Journ. Physiol., lxvi, 356.
Macleod and Knapp. 1919. Amer. Journ. Physiol., xlvii, 189.
Marshall. 1922. Journ. Biol. Chem., li, 3.
Meyer. 1924. Zeitschr. ges. exp. Med., xli, 52.
Munford and Hubbard. 1926. Journ. Amer. Med. Assoc., clxxxvii, 922.
Myers. 1922. Journ. Biol. Chem., liv, 675.
Myers and Booher. 1924. Journ. Biol. Chem., lix, 699.
Myers and Muntwyler. 1928. Journ. Biol. Chem., lxxviii, 225.
Pavlov. 1929. Conditioned reflexes.
Pawlow. 1910. Trans. by Thompson. The work of the digestive glands.
Peters and Van Slyke. 1931. Quantitative clinical chemistry, Vol. 1, 976, 979.
1932. Quantitative clinical chemistry, Vol. 2, 647.

QUINCKE. 1884. Zeitschr. klin. Med. vii, Suppl. 22.
RANNENBURG. 1926. Pflügler's, Arch., ccxii, 601.
ROBERTS. 1859. Edin. Med. Journ., v, 817, 906.
ROBINSON, SHEARN AND DAVIES. 1924. Quoted by KAYE, 1929.
RÜDEL. 1892. Arch. exp. Path. and Pharm., xxx, 441.
SCHULTEN. 1928. Münch. med. Wochenschr., xxi, 898.
SHERMAN AND GETTLER. 1912. Journ. Biol. Chem., xi, 323.
SIMPSON. 1924. Journ. Biol. Chem., lix, 107.
1926. Journ. Biol. Chem., lxvii, 505.
1929. Journ. Biol. Chem., lxxxiv, 393, 413.
SMALL. 1928. Glasgow Med. Journ., cx, 33.
STILL. 1931. Physiol. Rev., xi, 328.
TALBERT. 1919. Amer. Journ. Physiol., xlix, 127.
1920. Amer. Journ. Physiol., l, 579.
TALBERT ET AL. 1931. Amer. Journ. Physiol., xci, 426.
TAYLOR. 1931. Verbal communication at Physiol. Soc., June 10.
1932. Personal communication.
VAN SLYKE AND PALMER. 1920. Journ. Biol. Chem., xli, 567.
VAN SLYKE, STILLMAN AND CULLEN. 1917. Journ. Biol. Chem., xxx, 401.
WATSON. 1925. Can. Med. Assoc. Journ., xv, 32.
WILSON. 1923. Physiol. Rev., iii, 295.
WORDELL. 1924. Zeitschr. f. d. ges. exp. Med., xlii, 389.

THE PHYSIOLOGY OF HEARING: THE NATURE OF RESPONSE IN THE COCHLEA

ERNEST GLEN WEVER

Department of Psychology, Princeton University

Recent progress in the understanding of cochlear function has been achieved by two fairly distinct methods, first, by elaboration of the classical theories to account for an increasingly detailed knowledge of auditory phenomena, and second, by the development of nerve physiology and realization of the limitations which nerve transmission processes impose upon the essential activities of the end-organ. This review, therefore, after a brief outline of prevailing theories, will begin with the nature of cochlear response in relation to the facts of hearing, and will then turn to the recent advances in the physiology of nerve in their bearing upon acoustic problems. Finally, these two lines of development will be brought together in an effort to outline the trends of current thought in auditory theory.

THEORIES OF HEARING. All modern conceptions of sound reception in the cochlea are based upon the anatomical evidence that the true sensory endings for hearing are the hair cells of the organ of Corti lying upon the basilar membrane. The remaining parts serve an accessory function of communicating sound pressures to these sensory cells.

Among the accessory structures the basilar membrane plays the most prominent rôle. Since it separates the scala vestibuli from the scala tympani, it is the obvious pathway for discharge of pressures through the cochlea. The pressure discharge from stapes to round window causes transverse displacements of the membrane, which lead in turn to stimulation of the sensory cells and of the nerve terminations which these cells contain. Thus far the theories are agreed.

But consideration further of the behavior of the basilar membrane system in translating sound vibrations into nerve responses leads into a field of varied speculation. Traditionally, the many theories in this field have been divided into two rival classes: 1, those that assume the basilar membrane (or related structures) to be differentiated in response properties so that particular tones give rise to vibratory movements

in characteristic regions, and 2, those that assume an undifferentiated membrane that responds alike to vibrations of any frequency. Theories of the first class are called *place theories*, because pitch is correlated with the response of particular regions of the membrane, and hence with the excitation of specific fibers of the auditory nerve. Theories of the second class are called *frequency theories*, because pitch is more directly accounted for by the transmission to the auditory nerve fibers, and ultimately to the higher centers, of impulses faithfully representing the frequency of the mechanical vibrations.

Of the several types of theory included in the above classes, three only will be considered here, for only these have sought to explain comprehensively the facts and phenomena of hearing, and thus have figured prominently in recent discussions. They are 1, the resonance type of place theory, known also in honor of its founder as the Helmholtz theory; 2, the "telephone" or simple frequency theory, formulated by Rutherford and developed recently by Boring, and 3, the analysis type of frequency theory, or "hydraulic" theory of Max F. Meyer.

The resonance theory.[1] This theory invokes the principle of sympathetic vibration, or resonance, as a basis for frequency selection in the cochlea. As is well known, a resonator will respond most actively to frequencies approaching that to which it is tuned; and if a series of resonators is arranged with progressively varying natural frequencies, a given vibration applied to the system will produce a response characterized by its position along the series. In the case of the ear, it is assumed that the basilar membrane is graduated in resonance properties, so that any given tone of the audible scale has its characteristic region of response. When a particular part of the membrane is excited, the nerve fibers supplying it are stimulated, and transmit impulses which lead to the perception of a tone of particular pitch.

When two or more frequencies are simultaneously presented, the response is distributed over the membrane accordingly, each frequency with its own place and amplitude of movement. A complex sound, therefore, is analysed in a spatial sense, its pattern of sound pressures being transformed into a topographical pattern of vibratory response on the membrane. The pattern on the membrane is reflected in the pattern of excitation of auditory nerve fibers, and thereby is presented to the higher cortical centers.

The result of such analysis is conceived as roughly that achieved by

[1] For current formulations of this theory, see Wilkinson and Gray (32), Hartridge and Banister (18), and Fletcher (10).

mechanical or mathematical methods, through which a compound wave is resolved into a harmonic series of sine (or cosine) components of particular frequencies and amplitudes.

Intensity in this theory is represented by both the amplitude and the spread of response on the membrane. Hence it is represented in the nerve response, according to present neurological conceptions, in terms of the number of fibers acting and the frequency of impulses in these fibers.

The "telephone" theory.[2] According to the "telephone" or simple frequency theory, the basilar membrane is not differentiated in its response to the various tones of the scale, but any tone may set a part or the whole of the membrane in vibration. When the membrane is actuated by a tone of a given frequency, the hair cells, and through them the auditory nerve fibers, are excited at that same frequency. Hence the correlate of pitch in the nerve transmission process is frequency of impulses in the fibers concerned.

The number of fibers transmitting this frequency is a function of intensity. The greater the intensity of the sound, the greater will be the spread of response over the extent of the membrane, and hence the greater the number of nerve fibers excited.

Complex sounds, on this theory, are not peripherally analysed, but such analysis as experientially can be demonstrated is regarded as a function of the cortical centers.

Meyer's "hydraulic" theory. The type of frequency theory advanced by Max F. Meyer (21) differs fundamentally from the one just discussed in that by postulating particular properties and modes of action of the basilar membrane it provides for the peripheral analysis of complex sounds. According to this theory, the basilar membrane is laxly suspended across the bony canal between the two scalae, and is free to move a short distance up or down but no more. Hence, when pressure is exerted at the stapes, the adjacent parts of the membrane are displaced downward, but as the pressure increases they reach their limit of movement and other portions somewhat more remote must be included in the action. Thus the displacement spreads as the pressure rises, always beginning near the stapes and extending along the membrane toward the apical end of the cochlea. The distance of spread is dependent upon the amount of the pressure change, and to a minor extent also upon its rate of change. When the stapes reverses its movement and pulls out-

[2] The theory discussed here resembles that of Boring (5), but omits any consideration of *volume* as an extensive attribute of tones.

ward, a second displacement begins, also near the stapes, but in this case a displacement upward, which likewise spreads along the membrane. The second displacement may be regarded as cancelling the first, for at the end of the first impulse the membrane, because it is nearly inelastic, remains essentially at rest until another impulse of contrary direction alters its position.

The greater the number of changes of pressure, or, in a simple wave, the higher the frequency, the greater will be the number of displacements of the basilar membrane and its related structures, and hence finally, on this theory, the number of impulses in the nerve fibers supplying the region concerned.

When two or more periodicities are combined in a compound wave, the pressure changes exerted by the stapes produce a complex series of displacements which may be regarded as made up of several sets or patterns of displacements differentiated by both rate and distance of spread; since the distance as well as the rate characterizes these separate patterns it follows that the frequency components of the complex wave are topographically distributed on the membrane. Consequently the frequency of excitation of sensory elements will vary at different regions of the membrane, and a number of distinct impulse frequencies will pass through separate groups of fibers of the auditory nerve. The intensity of each separate frequency is represented by the length of basilar membrane activated at the rate in question, and thus in the number of nerve fibers stimulated at this rate.[3]

AUDITORY PHENOMENA IN RELATION TO COCHLEAR FUNCTION. The evidence upon which the above theories must be judged consists of a, physical and anatomical evidence; b, evidence from psychological phenomena, and c, evidence from human and animal pathology.

a. *Physical and Anatomical Evidence.* As the basilar membrane is the central feature of practically every theory of cochlear action, this structure has been studied with much care. Histological examination shows the form of the membrane to vary considerably in its upward course within the cochlear canal. Its width is least at the basal end, and increases in a somewhat irregular manner to a maximum a short distance from the apex. The variation in width, according to recent measurements, is about three-fold (Guild, 13), which is considerably below the twelve-fold variation formerly accepted. A second feature is that

[3] A second form of analysis-frequency theory was developed by Wrightson, but as the chief feature in which it differs from the simple frequency theory has met with severe criticism the theory is not considered separately.

the cell masses lying on the membrane show variation in bulk, increasing from basal to apical ends. A third feature is that the spiral ligament is larger near the base and diminishes in size toward the apex.

These structural variations have been used as a basis for the differentiation of response assumed by the resonance theory. The increase in length of fibers from base to apex is regarded as giving a primary gradation of tuning from high to low. From the structure of the spiral ligament it is inferred that tension is exerted upon the transverse fibers of the membrane, a tension that is great near the base where the ligament is larger, and falls off rapidly toward the apex where the ligament is small. The variation of cell masses is considered a progressive loading of the fibers, less near the base, greater toward the apex. It is also suggested by Wilkinson (32) that further loading is afforded by the columns of fluid in the two scalae, measured from oval and round windows to the fiber in question, which must be set in motion when the fiber is moved; such columns obviously increase in length with the distance of the fiber from the base of the cochlea. All these factors, length, tension, and mass, vary in the same effective sense, so as to give the fibers at the base the highest natural frequencies and those toward the apex progressively lower frequencies. And it is argued that the variation in these physical characteristics is sufficient to give a graduation of natural frequencies from the lowest to the highest audible tones.

Outstanding criticisms of the resonance hypothesis are, first, that differentiation by tension either does not exist or is insufficient to give a variation of response of the necessary range, and, second, that the fibers of the basilar membrane are so bound by longitudinal connections that they could not possibly act as individual resonators.

1. According to the most favorable estimates of differentiation by length and mass, the factor tension must vary by about 5000-fold to account for the difference of tuning between lowest and highest audible tones. Many are not willing to concede the possibility of variation of tension of such order.

An important modification of the resonance hypothesis denies the necessity of so wide a variation of tuning as that just considered. As is well known, a mechanical system may be forced to respond, though at low amplitudes, to frequencies far outside its natural frequency. The tuning of the ear may be limited to the middle range of tones, and the extreme tones may be operative only through forced response of the end resonators. The rapid decrease in acuity for the extreme high and low tones suggests that these are beyond the natural tuning of the system,

but nevertheless when imposed in sufficient strength are able to actuate the terminal regions of the membrane (11).

2. The concept of individually acting resonators in the cochlea has been wholly abandoned in modern theory. Helmholtz himself recognized that the fibers of the basilar membrane are not free, but are bound by the neighboring tissues, and he therefore supposed a narrow band of the membrane to be actuated by a single tone. The modern tendency has been to expand more and more the concept of spread of response on the membrane, until most exponents of the resonance hypothesis conceive of relatively large sections of the membrane as involved in the response to any given frequency. It has been recognized that the frictional resistance incurred by the moving tissues must be very large, and the damping of resonance consequently high, which by physical principles leads to a broadening of tuning.

Moreover, histological studies (19, 22) have shown considerable diffuseness of the nerve supply of the organ of Corti. Many of the nerve fibers are branched near their terminations in the sensory cells, and thus one fiber may serve a group of cells. Therefore if mechanical response could indeed be limited to individual basilar membrane fibers, or even narrow bands of the membrane, it could have no correspondingly restricted representation in the nerve transmission process.

The simple telephone theory rests upon no criticial response characteristics of the membrane, and supposes only that the spread of response is greater the higher the intensity, a principle suggested first by Meyer, but now admitted by all theories.

The Meyer theory, on the other hand, requires that the membrane be nearly inelastic, that it be not stretched but supported loosely, and that its movements be completely damped.

b. *Evidence from Psychological Phenomena. Pitch discrimination.* One of the foremost tasks of any auditory theory is to explain the basis for pitch discrimination. This capacity is most delicate for tones of the middle range, and falls off considerably for higher and lower tones.

The frequency theories represent the frequency as transmitted by way of the auditory nerve fibers to the higher centers, and hence relegate all pitch discrimination to these centers.

Obviously, the simplest type of explanation of pitch discrimination on a resonance theory would be to suppose a separate resonator for each discriminable tone. The theory in its earlier form followed this line of thinking. The necessary number of resonators was computed on the basis of the magnitudes of the difference limens at various regions of the

scale. Helmholtz calculated that 4500 resonators would suffice. More complete data on pitch discrimination have led to the recent estimate of 2000 discriminable tones (10). Such calculations, however, are based upon an assumption, not always expressed, that the sensation of pitch proceeds along the scale by discrete steps or quanta. Experience indicates, on the contrary, that if the tones 1000~ and 1003~ are just discriminable, so also are the tones 1001~ and 1004~. In general, at a given region of pitch we may choose the first tone where we will, and the size of the relative difference limen remains the same. If pitch changed by discrete steps we should expect to find critical points where a small change of stimulus produced a jump from one tone to the next adjoining. Thus if a critical point existed between 1001 and 1002~, this difference should as readily be distinguished as that between 1000 and 1003~. Such critical points have never been discovered in pitch, or indeed in any field of sensation.

Helmholtz in a sense anticipated a solution of the quantum difficulty in his theory without perceiving the difficulty itself, for in another connection he suggested that frequencies intermediate between those proper to an adjacent pair of Corti's arches may be differentiated by setting both these arches, and perhaps other nearby ones also, into forced vibrations with amplitudes depending, for each resonator, upon the degree of mistuning. The pitch then is determined by appreciation of the complex of relative amplitudes of the several resonators in action. More recent versions of the resonance hypothesis, in admitting a far higher degree of spread of response over the membrane, complicate this picture still further.

When on the resonance theory it is admitted that the response to a single tone is widely distributed over the membrane, it becomes a serious problem how any precise pitch perception is possible, and especially how pitch remains constant (or practically so) in the presence of changing intensity and consequently varying breadth of the response band on the membrane. An attempt to solve this problem was made by Gray (32) in what is called the principle of maximum stimulation. This principle states that the pitch is perceived as that proper to the point on the membrane where the amplitude of vibration is a maximum. In terms of nerve response, this means that the pitch is determined by the nerve fiber (or fibers) in which the rate of impulses is greatest.

Loudness discrimination. The capacity of perceiving differences of loudness varies with both frequency and loudness level. But for fairly loud tones the size of the relative difference threshold is practically con-

stant, which means that within limits the Weber-Fechner law holds for tonal intensity.

The problem of sensory quanta has been raised for loudness in much the same way as for pitch. But here also any experimental indications of quanta are lacking.

The resonance theory avoids any assumption of loudness quanta, since one of the determining factors, rate of nerve impulses, is continuously variable. The frequency theories, however, depend upon number of nerve fibers only as the intensive correlate, and since such fibers must be added as discrete units, this theory has been considered as committed to a quantum assumption for loudness. But it is doubtful if such a commitment is important practically, as the number of auditory nerve fibers is sufficient to account for steps considerably smaller than the difference limen, and quantum steps, if such existed, would be so small in comparison with the difference limen that they hardly could be revealed in experience.

The appreciation of phase. If a resonator, such as a tuning fork, is being driven by sound waves from some external source, and the phase of the incident waves is suddenly changed, the vibration of the fork will slowly die down to zero, and then as slowly rise again in agreement with the new phase condition. It was formerly believed that if resonance occurs in the ear something similar must happen: that if a tone were abruptly altered in phase we should experience a fluctuation in the loudness of the sound. Hartridge (17), on producing a phase shift of 180° in a tone, reported such a fluctuation. Békésy (3), however, using better technical methods, failed to discover any such effect.

Békésy's result is not incompatible with the modern form of resonance theory, since in the presence of high damping the response of structures in the cochlea should follow fairly closely even abrupt changes of phase.

Numerous experiments, dating from Helmholtz' time to the present, agree in showing that the relative phase obtaining between the several component frequencies of a complex wave is a minor factor in determining the quality of the perceived sound. Under some conditions, as experiments show, an observer can detect slight changes in sounds as a result of phase alterations. But in general it appears that phase is not very important in uni-aural hearing.

It may be argued that any appreciation of phase relations in a complex sound is incompatible with a resonance theory, since that theory assumes no strict temporal relation between stimulus and nerve response. Yet an indirect appreciation of phase conditions is possible on such a theory

if on the basilar membrane there is overlapping of the bands of response of different components. Such overlapping would give rise, in the regions common to two (or more) frequencies, to a compound vibration whose form would vary with the phase relations. And if, as seems likely, the form of distribution of response on the basilar membrane is a factor in the determination of loudness, that is, if nerve excitation is not a linear function of membrane displacement, a change of phase might under favorable conditions be made manifest in perception.

According to the frequency theories, perceptual changes due to phase alterations would likewise be explained, since these theories also assume overlapping of frequencies in their response on the basilar membrane.

Beats, combination tones, and masking. When two tones are conducted simultaneously to the same ear, three types of phenomenon may result. If the two tones differ but little in frequency, and are on something near the same level of intensity, there occur the characteristic surgings known as beats, whose rate increases with the frequency difference of the tones. If the frequency difference becomes sufficiently great, the result is a third tone of rather rough quality, called a combination tone, in addition to the two primaries. If the intensity of one of the tones is considerably above that of the other, the phenomenon of masking enters, and only the stronger tone is heard. That these phenomena are to be accounted for peripherally is strongly suggested by the fact that they do not occur when the two tones are led to separate ears.

Beats are explained on the resonance theory by supposing that the two generating tones involve somewhat overlapping regions of the basilar membrane. In such case the movement of the region common to the two tones may be represented as the resultant of the primary movements. When the primaries are in phase the upward and downward displacements of the membrane will be a maximum; when the primaries are in contrary phase these movements will be a minimum. Hence the part of the membrane in question will undergo periodic changes in amplitude corresponding to the rate of change of phase relations, which obviously is determined by the difference in frequency of the primary tones.

On this theory, the maximum number of perceptible beats should indicate at least the lower limit of spread of the generators on the basilar membrane. This maximum varies at different regions of the scale, being about 80 at 500~ and increasing to about 250 at 4000~ (28). It would therefore appear that the tone 4000~, for example, excites, to an extent sufficient to produce beats, resonators of natural frequencies

ranging from 3875 to 4125~; the total range might easily be greater than this but it could not be less.

According to this account, the maximum number of perceptible beats should increase with the absolute intensity of the generating tones. Beyond casual observations that beats in general are more noticeable when the generators are loud, there seems to be no evidence on this point.

Masking data, interpreted on the resonance theory, indicate indeed that the range of response is considerably greater than that shown by the maximum number of perceptible beats. A given tone alters the threshold of other tones over a wide band of frequencies both above and below it, though to an extent that decreases rapidly with the frequency interval (26). The size of the bands is a function of the intensity of the masking tone, but for a faint intensity, such as 20 db., a tone of 1200~ noticeably affects the perceptibility of all frequencies from about 600 to about 2000~. For tones of equal loudness, the absolute range of the masking effect increases with the frequency, and, at least for weak tones, at a rate closely comparable with the rate of increase in the maximum number of perceptible beats. If loud tones are used, the masking is greater in degree and in range, especially for tones higher in frequency than the masking tone. This dissymmetry of the masking curve for loud tones is explained as the result of the production of harmonics through mechanical distortion in the middle ear.

The explanation of masking is that if a given region of the membrane is already activated strongly by a given (masking) tone, a second tone will be ineffective at relatively low intensities because the increment of amplitude added by it at any given point is inadequate to stimulate any more sensory endings than are already in action there; and the second tone will remain inaudible until its intensity is raised to the extent that additional endings are affected either in some part of the region already considered or, by increase in the range of response, in adjoining regions. An increased stimulation of sensory endings will of course lead to the excitation of a greater number of nerve fibers, or to an increase in the rate of response of fibers already acting.

If this explanation of masking holds, and particularly if raising the second tone to perceptibility is mainly effected by increasing its range of resonance, then it should follow that the pitch of the second tone is altered in a direction away from the pitch of the masking tone, as obviously its movements on the membrane will be more effective at points distant from the region of the masking tone. Experimental data on this point have not been obtained.

The explanation of beats by the frequency theories is essentially in terms of the varying amplitude of the compound wave, and hence the varying amount of spread on the basilar membrane.

Meyer in his theory has explained the masking of high by low tones in terms of overlapping on the membrane. In the theory, low tones extend farther along the membrane than do high tones of comparable loudness, because inertia effects make spread to some extent dependent upon frequency. Masking in the direction stated is the more prominent form, but masking of low by high tones takes place also, and to be fully adequate the theory should be expanded to explain both.

Meyer's theory accounts for combination tones directly in terms of the normal action assumed for the cochlea in the analysis of complex waves. Analysis on this theory is not of the simple harmonic or Fourier type, but a number of periodicities result which were not mathematically present in the stapedial movements. Among these are the combination tones. A disturbing feature of the theory, however, is that the analysis yields other periodicities also about whose experiential existence there is some doubt. The periodicities are often at theoretical intensities that seem at variance with experimental data, though it must be admitted that data on this subject are few and somewhat uncertain. In many instances the rates of the impulses are not regular, though over a short period of time they may represent roughly a given frequency. Gaps often occur in a series of impulses, but this need not be regarded as a flaw in the theory, since it has been found that the omission of pulses in a sound wave does not seriously interfere with pitch perception.

The resonance theory has no explanation of the origin of combination tones in the cochlea itself, but assumes such tones to arise through distortion in the middle ear. If the movements of the stapes are not strictly representative of the pressure variations of the incident sound, but instead are sensibly proportional, say, to the square of the pressure, it follows from physical principles that the frequencies represented by the combination tones, as well as the primaries, will be presented to the cochlear system.

Auditory fatigue. Experiments on fatigue of the ear show considerable discordance of results. The recent work of Békésy (4) indicates that exposure to a given tone causes a temporary reduction of sensitivity to tones over a broad range on either side of the fatiguing tone. The fatigue effect is a maximum at the frequency of the fatiguing tone, and decreases rapidly with the distance from that tone. He reported also that tones lying above the fatiguing frequency are raised in pitch, while

tones below are lowered in pitch. Weinberg and Allen (27) found that the persistence of tonal sensation is altered for all tones in the region of a fatiguing tone, though in their experiments the range of the effect was comparatively small.

These phenomena are explained by the resonance hypothesis as a temporary impairment of a particular region of the basilar membrane or, perhaps more likely, of the sensory or nervous cells which serve that region. No explanation seems available in terms of the frequency theories.

The binaural localization of sound. The direction of sounds is appreciated by means of differences of intensity, time, and phase at the two ears. A sound is localized toward that ear where, within prescribed limits, the excitation is greater, or sooner, or of "leading" phase. The factors of localization by intensity and time are not crucial to theories of hearing, as they can be accounted for by any theory in terms of integration of the neural excitation processes from the two ears, provided only that the neural processes represent with fair accuracy the temporal and intensive properties of the incident sounds.

Localization in terms of phase, however, is of great theoretical significance. Plainly, if a binaural phase difference is to be appreciated, there must be some representation in the neural response of each ear of the phases of the incident tones. On a frequency theory such representation is given through the direct translation of frequencies into nerve impulses. On a resonance theory, however, phase has no obvious representation.

The attempt has been made to account for phase on the resonance theory by reducing relative phase to relative time on the assumption that nerve excitation occurs at some particular phase in the wave of stimulation. Such an explanation might serve if the appreciation of phase differences is limited to very low tones, but obviously an extension of the theory to the higher frequencies would make it necessary to assume in the resonance theory essentially as close a temporal correspondence between basilar membrane movement and nerve fiber excitation as is postulated in the frequency theories.

The problem of the frequency limits within which binaural phase differences are effective is yet unsettled. There is no question that low tones are relatively easy to localize in terms of phase, and that localization becomes progressively more difficult as the frequency is raised. Most investigators report that localization is uncertain or impossible for frequencies above 1000 to 2000~. Halverson (16), however, though agreeing that marked shifts of localization are limited to low tones, found

slight movements for all tones up to 11,000 to 13,000~, depending upon the subject.

Halverson's experiments, however, like many others in this field, are subject to criticism on the basis of the method used for varying phase. He used binaural tubes (similar to stethoscope binaurals) to conduct the sound from a single source to the two ears, and varied the phase relation by changing the length of one of the tubes. The criticism of this method is based upon the fact that variation of the length of a tube varies its resonance characteristics, and thus the intensity at its orifice. Hence intensive as well as phase differences are presented to the ears, and the resulting judgments of localization are a function of this complex rather than of phase alone. It is very desirable that the upper limit of frequency for which any degree of phase localization is possible be carefully investigated in a considerable number of subjects by other methods which avoid the criticism just mentioned.

c. *Evidence from Pathology.* Though the minute size and general inaccessibility of the cochlea have so far restricted direct examination of its functioning, considerable knowledge has been gained from investigations of changes in hearing as correlated with pathological conditions either fortuitous or artificially produced. Artificial lesions are of two kinds, those produced by mechanical destruction of local parts of the cochlea, and those caused by prolonged stimulation of the ear by intense sounds.

Experiments on stimulation deafness. A large number of experiments have dealt with the effects of long-continued stimulation on the histological structure of the cochlea.[4] The results of animal experiments of this kind show modifications of the histological character of intracochlear structures and in some instances also of nervous tissues contained within the modiolus. The modifications vary greatly in different cases, depending in part, it seems, upon the intensity and duration of stimulation. Changes in the structures lying on the basilar membrane are most prominent; these vary from slight alterations in the form of the hair cells to complete disappearance of the organ of Corti. The nerve cells and ganglion cells may also be affected, first undergoing alteration in appearance, and finally suffering partial or complete atrophy. These structural changes when produced by tonal stimuli are typically restricted to particular regions of the cochlea, those due to high tones in the basal region, and those due to low tones toward the apex. The extent of the injury

[4] For a summary of experiments in this field, see von Eicken (9).

is widespread, the regions ranging typically from an eighth to a fourth of the whole length of the cochlea. In some of the experiments attempts have been made to ascertain the nature of the changes in auditory ability in consequence of the stimulation, but these efforts have been limited by the general difficulties of auditory measurements on animals. They therefore have been restricted practically to observations of general bodily movements and of the pinna reflex as correlated with sounds. Results of such testing methods are not reliable, and the experiments should be repeated with tests by the conditioned response technique in common use in psychological studies of animal sensitivity.

Experiments on cochlear destruction. The surgical injury of limited portions of the cochlea in animals has been attempted several times, with subsequent efforts to ascertain the resulting modifications of hearing. Like the experiments just mentioned, these also have suffered from the lack of adequate testing methods. Corradi and Baginsky reported that destruction of the apex of the cochlea impaired hearing for low tones, but Stepanow in a similar experiment found hearing to remain for both high and low frequencies. The experiments depended upon general observations of behavior for diagnosis of auditory ability. Wittmaack (33) in a recent study destroyed all but part of the basal turn of the cochlea of a dog trained by the conditioned response method, and found at first response only to high tones, and after some time response to high and middle tones down to 340~, but not to lower tones. This type of experiment plainly should be repeated with the use of several animals and variation of the locus and extent of injury.

The most thorough of the studies of cochlear destruction, though impaired by the use of the pinna reflex for diagnosis of hearing ability, is the one by Held and Kleinknecht (20) on the guinea pig. These investigators sought to alter the tension of the suspensory ligament in a given region by carefully boring into it with a fine needle. The results of injury to the basal turn showed impairment of hearing for high tones within a limited range. Attempts to injure apical regions were unsatisfactory for the reason, as they believe, that the cochlea there is too fragile to withstand the operation and is harmed in a general and unpredictable way.

Results from human pathology. A few studies have been made of human temporal bones in cases where tests of hearing had been made preceding death. Such studies clearly are of great value in the problem of the relation between structural conditions and the functioning of the ear. But due to the manifold character both of anatomical variations

and of auditory impairments, a statistical treatment of large numbers of cases is necessary. An experiment satisfying this requirement is in progress at the Johns Hopkins Otological Research Laboratory, preliminary reports of which are already available (14, 15). In this experiment, routine hearing tests are made on numbers of hospital cases, and when autopsy material is available the temporal bones are studied. Two types of result are indicated: 1, a correlation between local degenerations of the organ of Corti and regional losses of tone perception, and 2, a like correlation between local atrophy of auditory nerve fibers or ganglion cells and regional losses of hearing.

The results of the above three types of investigation are well in agreement in indicating a relation between tonal frequency and cochlear region. The relation appears to be general rather than specific, in the sense that any tone activates a fairly wide band of the membrane. But the location of this band over the linear extent of the membrane is a function of frequency. High tones are localized toward the base, and low tones somewhat less definitely toward the apex.

These results seem generally favorable to a place hypothesis such as the modern form of the resonance theory. Results like those of Wittmaack on the dog, in which reaction to all but the lowest tones remained with only the basal half-turn of the cochlea, are explainable on this theory if it can be assumed that the high tones, though audible to the animal, were not discriminable. All high tones, by the principle of maximum stimulation, should in this case be of the same pitch or, more likely, of indefinite pitch.

A complete explanation of the above results in terms of the frequency theories has not been offered. Attempts have been made to account for the stimulation deafness results, such as those of Yoshii which are better known, in terms of the frequency theory.

1. Meyer suggested that, since Yoshii used tones of equal intensities, the various frequencies should have affected different regions of the membrane because on the "hydraulic" theory the distance traveled up the membrane is in part an inverse function of stapedial velocity, and low tones should extend farther than high tones of the same strength. But the weakness of this explanation is obvious when it is considered that while Yoshii for his various whistle tones did use air pressures at least approximately the same (4 and 5 mm. of mercury), it is hardly safe to say that the intensities as received upon the basilar membrane were anything near equal, as the mechanical impedance of the ear most assuredly does not remain constant over the four and a half octaves (from g to c^5) involved in the experiment.

2. Boring, in an attempt to account for the same results in terms of the simple frequency theory, assumed that the lower tones were the strongest, and hence that they would extend the greatest distances up the cochlea. This explanation is reasonable at least, since the impedance of the ear increases at high frequencies, and therefore, if Yoshii's tones were about equal to begin with, they would be graded in intensity at the basilar membrane in the direction required.

But this explanation in terms of relative intensity grows doubtful in contemplation of the assembled results of the many other investigations of stimulation deafness. It would seem very remarkable that, in general, the high tones should be weak and the low tones strong, despite a great variety of experimental conditions.

It is difficult to see how, on the frequency theories, an analogous explanation in terms of intensity can be applied to the results of cochlear destruction and the studies of human pathology.

Further Anomalies of Hearing. Tonal gaps. Tests of hearing sometimes reveal a marked loss of sensitivity over a limited region of the tonal scale; such a regional defect is usually called a tonal gap. The rate of decrease of sensitivity at the borders of such a gap is usually of the order found at the upper and lower limits of the normal auditory range. It was formerly believed that sensitivity to frequencies within a tonal gap was completely absent, but it has been shown that such a gap tends to close when explored with tones of high intensity, and some investigators report that gaps close completely when sufficiently loud tones are used. More data are needed on the pitch of tones heard within a tonal gap as compared with the pitch perceived by the other (normal) ear, and also on the magnitude of the difference threshold for pitch in the gap and especially on its borders. Such measurements as have been made on persons suffering from serious impairments of hearing have shown that for tones that they are able to hear the difference thresholds for both loudness and pitch are not significantly different from normal.

Tonal gaps are readily explained on the resonance type of theory as due to a local lesion of the basilar membrane or of its nervous elements.

Tinnitus. Internal sounds, or "ringing in the ear," form a fairly familiar experience. Fortunately, in all but a few cases, such phenomena are temporary; when continuous they are recognized as pathological symptoms. Tinnitus occurs in many forms. Some are dull or noisy in character, and are usually ascribed to circulatory disorders or to mechanical defects in the conductive system of the middle ear. Others are tone-like, and often of high pitch; these are believed to have their origin in the

inner ear or its nervous connections. Tonal tinnitus is usually of fairly definite pitch. Exploration of the region of the tinnitus with objective tones reveals various phenomena of beats and masking, showing that the tone of internal origin operates in much the same way as would an objective tone of the same apparent frequency (25).

Tinnitus, in some of its forms at least, is explained by the resonance theory as due to a regional excitation of sensory or nervous elements such as might arise through the chemical products of a local lesion.

Diplacusis. An unusual type of acoustic anomaly is one in which a tone is perceived as of different pitch in the two ears. There are many forms, but the most frequent case is one in which tones within a limited region of the scale are heard as noises, or as tones with a prominent noisy quality. Another form is one in which certain tones are heard as raised or lowered in pitch as compared with their perception in the other (normal) ear.

The "noise" type of diplacusis is probably best accounted for as due to serious distortion in the conductive mechanism. The type in which displacement of pitch occurs, however, requires different explanation. On the resonance theory, this phenomenon is regarded as resulting from a change in the resonance characteristics of a part of the basilar membrane such that certain frequencies affect regions slightly different from those formerly affected. Since pitch is associated with particular region, and hence with particular nervous elements, such displacement of response would lead to a shift in the pitch as perceived by the affected ear.

REVIEW OF THE ABOVE EVIDENCE. The chief developments of the above discussion may now be briefly summarized.

The anatomical facts, so far as they are known, do not especially favor any of the theories treated. The particular mechanical characteristics assumed by the theories differ greatly, but crucial physical and physiological data are lacking. It is agreed that the damping in the cochlea is of a high order, with the result that the resonance theory no longer presents the simple picture of "a pitch to a fiber" as formerly, but conceives a single tone as spreading extensively over the membrane. The "resonance" in the theory is remote from that represented in the familiar analogies of piano and harp.

The primary qualities of pitch and loudness are about equally well accounted for by the theories. All the theories require complex differential and integrative processes in the higher centers. The frequency theories require central appreciation and differentiation of impulse rates for pitch, while the resonance theory requires appreciation of impulse

rates and their integration over several fibers for intensity. The frequency theories require appreciation of the extent of spread for intensity, while the resonance theory requires appreciation of the locus and form of spread, under the principle of maximum stimulation, for pitch. It is apparent that none of the theories can make any claim of advantage on a basis of simplicity.

The evidence from beats, masking, and fatigue, as well as that from pathology, appears definitely more favorable to the resonance type of theory. A spatial distribution of frequencies seems to form a more reasonable basis for the functional interaction of adjacent frequencies and the selective impairment of tonal sensitivity.

On the other hand, the fact of the appreciation of binaural phase differences definitely favors the frequency type of theory.

THE NATURE OF AUDITORY NERVE RESPONSE. It is apparent throughout the above discussion that the theories are forced constantly to consider the nature of sensory nerve processes. The resonance theory assumes specificity of nerve fibers as the correlate of pitch, and frequency of impulses in the fibers as the correlate of loudness. The frequency theories, on the contrary, assume frequency of nerve impulses for pitch, and number of fibers for loudness. Contemplation of these assumptions concerning the character of the nervous excitatory and transmission processes leads to a careful examination of such evidence as is now available from nerve physiology.

Beside the specific energies doctrine, whose acceptance or rejection differentiated the early theories in their explanation of pitch, and the all-or-nothing principle, which led to important revisions in their account of intensity, there stands a third neurological factor of crucial importance, that of refractory phase.

Refractory phase. The maximum frequency at which a nerve fiber can respond is primarily defined by the length of its absolute refractory phase, that is, the time that must elapse after a given impulse before the fiber has recovered and is again stimulable. Most measurements of absolute refractory phase in mammalian nerves have yielded values of the order of one or two milliseconds, and the smallest value so far reported is 0.58 ms. for the phrenic nerve of the dog (12). Moreover, the refractory period has been found to be increased if the nerve is maintained in rapid response.

An absolute refractory phase of 0.58 ms. would permit a maximum frequency of only about 1700 impulses per second, which is much too low to account for pitch perception on a frequency theory. To transmit

the tones of 20,000 or more at the upper limit of the scale an auditory nerve fiber would need to have, and continuously maintain, an absolute refractory phase as short as 0.05 ms. If the auditory nerve fibers have refractory properties comparable with those of the nerves so far measured, a serious obstacle is presented to frequency theories of hearing.

The resonance theory suffers no disadvantages from limitations imposed by refractory phase, as loudness rather than pitch is correlated with rate of firing, and a moderate variation in response rate is sufficient, along with the number of fibers acting, to account for all degrees of loudness.

In defense of the frequency theories it is held that the auditory fibers may possess properties significantly different from those of other nerves. An absolute refractory period of 0.05 ms., it is argued, is only about as much below the shortest period found (0.58 ms.), as this is below the periods of many other fibers.

The situation clearly calls for measurements of the response properties of the auditory nerve itself.

Experiments on auditory nerve impulses. A series of experiments on auditory nerve impulses was begun recently by Wever and Bray (29, 31) in an effort to obtain data bearing upon the fundamental processes of hearing. Their results, which have been verified and extended in the researches of a number of other investigators, notably those of Adrian, Bronk, and Phillips (1) and of Davis and Saul (7, 8), give information, at least of preliminary nature, on the correlate of pitch in the auditory nerve response.

The method of the Wever and Bray experiments consisted of observing from the auditory nerve or from acoustic tracts of the medulla and brain stem of a decerebrate animal the action currents resulting from auditory stimulation. After suitable amplification of the action currents, observations were made by means of a telephone receiver or, in some cases, by means of a cathode-ray oscillograph. The conditions made possible the observation of impulses over a wide range of frequencies.

The observations showed, for low frequencies at least, a correspondence between the frequency of sound and the number of impulses set up in the auditory nerve and higher auditory tracts of medulla and brain stem.

The upper limit of frequency for which this correspondence is true is yet to be determined. The observations on this point so far have been uncertain because of the technical difficulty of distinguishing between impulses originating in nervous conduction fibers and other electrical changes which have a sensory origin. Observations made at the audi-

tory nerve itself yield results for frequencies up to about 5000~, but in these results the presence of sensory effects has not been excluded.

The nature and origin of the sensory effects is undetermined also. They result from intracochlear processes set up by sound stimuli, but it is debatable whether they are produced incidentally by the movements of fluid or membranous structures, or whether they arise from the action of sensory or ganglion cells. The possibility of sensory or nervous origin seems the more likely, as the effects are reduced or abolished by such operations as anesthetizing, cooling, and polarizing with direct current, all of which should not seriously affect the mechanical operation of cochlear tisssues (1, 31).

The complicating influence of sensory effects can be avoided by suitable technical measures, which are most readily carried out in recording from tracts of the medulla and brain stem, since the sensory effects suffer rapid decrement with distance from the cochlea. By the use of small shielded electrodes and relatively low amplifications, results can be obtained from the central tracts which are free of the sensory effect. The results thus obtained by Wever and Bray and by Davis and Saul agree in showing auditory nerve responses synchronized with the exciting frequencies up to about 1000 cycles. Wever and Bray are of the opinion that nerve responses of yet higher frequencies are present, though the evidence on this point is insufficient for a final decision. But, as Wever and Bray have pointed out, it is not likely that individual nerve fibers respond at the high rates found. High frequencies they consider to be a composite result of the "staggered" firing of a number of fibers; this is the *volley hypothesis*, discussed below.

The representation of intensity in the auditory nerve response has been only slightly investigated. Wever and Bray (31) found that the strength of the nervous response increases with intensity of stimulation. Preliminary measurements indicate further that for a given increase in sound intensity the amount of increase of nerve voltage is less the higher the level of intensity of the applied sound. In other words, something of the order of the Weber-Fechner law is operating between sound stimulus and nerve response.

The Volley Hypothesis. It is possible to formulate a principle by which a relatively high frequency can be represented in the response of a nerve as a whole, without assuming for the fibers which it contains any properties other than those which nerve fibers are known to possess. This, the volley principle, may be stated as follows:

A given nerve fiber, when stimulated by the action of the sensory cell

that it serves, responds, and then goes into refractory phase until again it is in condition to be excited. If the stimulus is periodic and continuous, somewhere in the relative refractory phase, just where depending upon the intensity of excitation, the fiber will again respond; and so continuously during the maintenance of the stimulus. If the excitatory processes mediated by the sensory cell are synchronized with the stimulus, the responses of the fiber will be synchronized also, though these responses may be considerably fewer in number than the pulses of stimulus. Hence for a given periodic stimulus and a constant rate of recovery, the fiber will respond with a regular series of impulses, which, though of lower rate than the stimulus, will always be dependent upon its frequency.

A second fiber, also excited by the same stimulus, will respond similarly, but if its rate of recovery varies slightly from that of the first fiber, its train of impulses, though strictly in synchronism with the stimulus, will be out of step with that of the first fiber. A large number of fibers, with various refractory periods, will respond with a total volley of impulses representing in the aggregate the whole stimulus frequency. A frequency of 20,000 per second could readily be represented by the combined response of 40 fibers each acting at a rate of 500 per second, provided that their times of entering the volley were arranged in just the proper way; in actual experience one would expect a frequency as high as 20,000 to require several hundreds of fibers acting at the rate mentioned.

It is evident from the nature of stimulability of nerve fibers during relative refractory phase that the greater the intensity of the excitation process the sooner in the relative refractory period will the fiber respond. Hence the greater the intensity of stimulation the more rapid will be the rate at which any given fiber enters the volley. Therefore it follows that, for any given bundle of fibers, the greater the strength of stimulus the greater the absolute number of nerve impulses transmitted by the bundle, but this without change in the frequency represented in the nerve as a whole.

This assumption for intensity is in agreement with Adrian's intensity-frequency principle, now well established in other sensory nerves, which states that the rate of response in a nerve fiber is a function of intensity of stimulation.

It is not necessary to the theory to assume constancy of refractory phase in the fibers; a change in refractory phase of a given fiber would alter its rate of entering the volley, but would not impair its synchronism with the exciting stimulus.

Nor is it necessary to assume that the response of a fiber corresponds to any absolutely fixed phase position in a stimulating wave, though it is assumed that the excitation of fibers by way of the basilar membrane occurs but once in a complete cycle, and is dependent in a general way upon the attainment of some given phase position. For lower frequencies of sound, at least, a small amount of variation of the temporal relation between nerve impulse and stimulus wave would not seriously impair the periodicity of the volley, but would merely round out the peaks in the response curve. Yet such variations might become significant if for high frequencies they approached in magnitude the periods of the volley impulses themselves.

It is apparent that the theory assumes transmission processes in the end-organ which make possible a close temporal relation between the external manifestations of stimulus and the immediate action of excitation of nerve fiber terminations. Whatever processes occur, mechanical, chemical, or electrochemical, must proceed at a rate rapid enough to prevent fusion of separate stimulus impulses. For a sense organ like the ear, it should be sufficient to assume that the sensory cells serve simply a mechanical function of communicating to the nerve terminations the motions of the stimulus. The strain exerted upon the naked ends of the axis cylinders, by altering the thickness of their surface layers, should lead to changes in membrane permeability and hence to excitation.

Since the formulation of the volley principle for audition, the principle has been invoked as an explanation of synchronous high-frequency discharges in nerves supplying pressure receptors in the skin. Adrian, Cattell and Hoagland (2, 6), on stimulating the frog's skin with an intermittent air blast, found impulses which followed the stimulus frequency up to about 400 per second. These results are most reasonably explained as a composite product of a number of nerve fibers acting on the volley principle.

Data so far obtained on auditory nerve responses are not crucial either for or against the volley hypothesis. The status of this hypothesis at present is determined by its service in explanation of the phenomena of hearing.

THE VOLLEY PRINCIPLE IN AUDITORY THEORY. *General considerations.* The outstanding advantage of the volley principle in auditory theory is that it permits a frequency theory of pitch without encountering any limitations of refractory phase in explaining the highest audible tones. Since on the basis of other considerations related above it is assumed that any given frequency spreads broadly over the basilar membrane and in-

volves a large number of nerve fibers, even the highest frequencies could be transmitted with at least fair fidelity.

A theory incorporating this principle therefore is in a position to explain localization in terms of binaural phase differences, since from each ear the frequency is transmitted in its particular phase relation. The increasing difficulty of phase localization with higher tones may be accounted for on the assumption of slight irregularities in transmission arising, as suggested above, from minor variations in the temporal relation between stimulus and nerve impulse; such irregularities naturally would be unimportant for low tones, but would grow increasingly serious at high frequencies.

In consequence of the operation of this principle several vibrations should be necessary in the case of the higher tones for the accurate determination of pitch, as after the initial impulse some time must elapse before sufficient fibers enter the volley to establish the frequency. The higher the frequency, the greater the number of vibrations required for pitch perception; and this is in agreement with present data on the minimal number of vibrations necessary to determine pitch.

The volley principle is able to account for the wide variations of both loudness and pitch without any assumption of sensory quanta. Pitch is represented as a continuum because it is a function of a continuous quantity, the volley-frequency. Intensity is continuously represented also because it is a function of total impulses per unit time, that is, of (composite) impulse rate, which likewise may vary without interruption. Since the theory includes Adrian's intensity-frequency principle, it gives a peripheral explanation of the Weber-Fechner law, from the form of the curve of impulse frequency as a function of intensity of stimulation.

The volley principle may be incorporated with any of the modern theories, but, in the writer's view, when combined with the resonance hypothesis it gives a composite theory which explains the phenomena of hearing in a more satisfactory manner than is possible for any other theory.

The Resonance-Volley Theory. It is a remarkable feature of the history of this subject that the resonance and frequency principles, though by no means inconsistent in themselves, have nevertheless continued through more than half a century as vigorously opposed hypotheses. That these are not incompatible in fact, but may readily be fused into a common, consistent theory, was pointed out by Rutherford (23) as long ago as 1898. But no one advanced such a theory until recently,

when formulations were made by Fletcher (11) and Troland (24), and developed by Wegel (25) and Wever and Bray (30). Fletcher proposed the volley principle in order to adapt the resonance theory to an explanation of phase localization.

Pitch, on the resonance-volley theory, has a two-fold representation, in terms of place on the basilar membrane and in terms of (composite) impulse frequency. How pitch may be determined by place has already been discussed; it is only necessary to recall that with great spread of response and hence many fibers in action any precise determination of pitch, as by the point of maximum stimulation, involves an appreciation of what fibers have the highest impulse rates, and hence assumes some kind of central differentiating process. The determination of pitch in terms of the volley frequency, on the other hand, involves an integrating process, since the several impulse series of many fibers must be combined.

If pitch is represented by both frequency and place, it becomes somewhat easier on a resonance theory to explain the pathological result that a high degree of deafness is accompanied by normal pitch discrimination for such tones as can be heard. It also becomes easier to account for Wittmaack's demonstration that a dog with but a fragment of a cochlea can hear a wide range of tones; and especially if it can be shown that such an animal retains the ability to discriminate pitch throughout that range.

Intensity on the resonance-volley theory, as in the resonance theory proper, is regarded as a function of the total number of nerve impulses reaching the centers, and thus is determined by both the number of fibers acting and their individual rates.

The more complex phenomena of analysis, beats, and masking, as well as pathological effects, are accounted for as in the simple resonance theory in terms of the spatial differentiation of response in the cochlea. These have been discussed earlier and require no elaboration here. In general it may be said that this theory combines the advantages of both frequency and resonance theories.

It will be appreciated that here is no simple theory, for even the elementary attributes of pitch and intensity require complex central processes. But no theory can be simple and pretend to a comprehensive explanation of the phenomena of hearing. Because the resonance-volley theory seems more comprehensive than any other, and at the same time is in harmony with our knowledge of nerve physiology, the writer regards it as the most reasonable theory of hearing at the present time. But

this view, as it is hardly necessary to say, is held subject to whatever evidences further experimental endeavors shall bring forth.

BIBLIOGRAPHY

(1) Adrian, E. D., D. W. Bronk and G. Phillips. The nervous origin of the Wever and Bray effect. Journ. Physiol., 1931., lxxiii, 2p.

(2) Adrian, E. D., McK. Cattell and H. Hoagland. Sensory discharges in single cutaneous nerve fibres. Journ. Physiol., 1931, lxxii, 377.

(3) Békésy, G. v. Zur Theorie des Hörens. Physik. Zeitschr., 1928, xxix, 799.

(4) Békésy, G. v. Zur Theorie des Hörens. Physik. Zeitschr., 1929, xxx, 115.

(5) Boring, E. G. Auditory theory with special reference to intensity, volume, and localization. Amer. Journ. Psychol., 1926, xxxvii, 157

(6) Cattell, McK. and H. Hoagland. Response of tactile receptors to intermittent stimulation. Journ. Physiol., 1931, lxxii, 392.

(7) Davis, H. and L. J. Saul. Action currents in the auditory tracts of the midbrain of the cat. Science, 1931, lxxiv, 205.

(8) Davis, H. and L. J. Saul. The frequency of impulses in the auditory pathways. Amer. Journ. Physiol., 1932, ci, 28.

(9) Eicken, C. v. Methoden zur experimentellen Schädigung des Gehörorgans durch Schalleinwirkung. Abderhalden's Handb. d. biol. Arbeitsmethoden, 1929, v, 771.

(10) Fletcher, H. Speech and hearing. 1929.

(11) Fletcher, H. A space-time theory of hearing. Journ. Acoust. Soc. America, 1930, i, 311.

(12) Gasser, H. S. and J. Erlanger. The nature of conduction of an impulse in the relatively refractory period. Amer. Journ. Physiol., 1925, lxxiii, 629.

(13) Guild, S. R. The width of the basilar membrane. Science, 1927, lxv, 67.

(14) Guild, S. R. Correlations of histologic observations and the acuity of hearing. Acta otolaryngol., 1932, xvii, 207.

(15) Guild, S. R., S. J. Crowe, C. C. Bunch and L. M. Polvogt. Correlations of differences in the density of innervation of the organ of Corti with differences in the acuity of hearing. Acta otolaryngol., 1931, xv, 269.

(16) Halverson, H. M. The upper limit of auditory localization. Amer. Journ. Psychol., 1927, xxxviii, 97.

(17) Hartridge, H. A vindication of the resonance hypothesis of audition, II. Brit. Journ. Psychol., 1921-22, xii, 142.

(18) Hartridge, H. and H. Banister. Hearing, in Foundations of experimental psychology, ed. by C. Murchison, 1929, 313.

(19) Held, H. Die Cochlea der Säuger und der Vögel, ihre Entwicklung und ihr Bau. Handbuch der normalen und pathologischen Physiologie, ed. by Bethe et al., xi; Receptionsorgane I, 1926, 503.

(20) Held, H. and F. Kleinknecht. Die locale Entspannung der Basilarmembran und ihre Hörlücken. Pflüger's Arch., 1927, ccxvi, 1.

(21) Meyer, M. F. The hydraulic principles governing the function of the cochlea. Journ. Gen. Psychol., 1928, i, 239. See references there cited.

(22) POLJAK, S. Ueber den allgemeinen Bauplan des Gehörsystems. Zeitschr. f. d. gesammt. Neurol. u. Psychiat., 1927, cx, 1.
(23) RUTHERFORD, W. On tone sensation with reference to the function of the cochlea. Brit. Med. Journ., 1898, ii, 353.
(24) TROLAND, L. T. The psychophysiology of auditory qualities and attributes. Journ. Gen Psychol., 1929, ii, 28.
(25) WEGEL, R. L. A study of tinnitus. Arch. otolaryngol., 1931, xiv, 158.
(26) WEGEL, R. L. AND C. E. LANE. The auditory masking of one pure tone by another. Phys. Rev., 1924, xxiii, 266.
(27) WEINBERG, M. AND F. ALLEN. On the effect of aural fatigue upon the critical frequency of pulsation of tones. Phil. Mag., 1924, xlvii, 126.
(28) WEVER, E. G. Beats and related phenomena resulting from the simultaneous sounding of two tones. Psychol. Rev., 1929, xxxvi, 402.
(29) WEVER, E. G. AND C. W. BRAY. Action currents in the auditory nerve in response to acoustical stimulation. Proc. Nat. Acad. Sci. (*Washington*), 1930, xvi, 344.
(30) WEVER, E. G. AND C. W. BRAY. Present possibilities for auditory theory. Psychol. Rev., 1930, xxxvii, 365.
(31) WEVER, E. G. AND C. W. BRAY. The nature of acoustic response: the relation between sound frequency and frequency of impulses in the auditory nerve. Journ. Exper. Psychol., 1930, xiii, 373.
(32) WILKINSON, G. AND A. A. GRAY. The mechanism of the cochlea, 1924.
(33) WITTMAACK, K. Demonstration zur Resonanztheorie. Acta otolaryngol., 1928, xii, 32.